The Dissertation Committee for Taegun Kwon Certifies that this is the approved version of the following dissertation:

Consequences, Repair, and Utilization of an Induced Double-Strand Break in the Chloroplast DNA of *Arabidopsis* and Tobacco

Committee:

David L Herrin, Supervisor

Stanley J. Roux

Robert K. Jansen

Enamul Huq

Arturo De Lozanne

Aravind Somanchi

Consequences, Repair, and Utilization of an Induced Double-Strand Break in the Chloroplast DNA of *Arabidopsis* and Tobacco

by

Taegun Kwon, B.S.; M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2011

This dissertation is dedicated to my father-in-law,

Jaehoon Oh (1943-2008),

whose sole unfulfilled wish was to caress and bless

his only beloved daughter

once and for all.

Acknowledgements

I would like to savor this special moment to think of the tremendous support I have received from a number of people for the past few years. I am sincerely grateful for such support. I thank Dr. Aravind Somanchi, who was my first professor to meet in the New World and could have been my committee chair, for his support during my brief stay at Auburn University and his aid in finding a new opportunity in Austin. He was certainly the one I would have most enjoyed to work with; I still believe he and I could have been a terrific "skinny scientists" duo.

My special thanks goes to Sunhee Oh, who has always been supportive despite whatever circumstances I have been in. I thank her, not just because she is my wife, but because she also has made a substantial contribution to my research. Whenever I became disheartened by the repeated unsatisfactory results, Sunny would carry out some of the experiments in my stead. Often, she was successful in bringing great fortune to my work. Notably, she was the first person to amplify double-strand-break repair products in *Arabidopsis* chloroplast.

I appreciate Dr. Obed W. Odom for the helpful discussions and joyful chats about nature, science, politics, culture, world, life and football. He was not only the best colleague in the lab but also a great guide for American culture and an excellent English teacher; however, my awkward English is not from him. I have also benefited from his good-enough German translation. He helped proofread this dissertation and provided ample constructive suggestions. Dr. Odom was also warmhearted enough to pose as a

Santa for my children every Christmas. His good spirit must have been inscribed deeply in the young souls.

I thank Liming Luo and Weihua Qiu for being the best lab mates and for sharing the critical moments in our lives. Weihua was the best person in the world to purify top quality I-*Cre*II enzyme, and generously allowed me to use her precious preparations, which were crucial for part of my research. Now she is in New York, which is one of the reasons why I am leaving this school. I also thank the amiable lady, Kristina Mahan, who welcomed my family to this city and made the Herrin lab a comfortable place to work five years ago. Brad Venghaus, Micah Cash and Dipesh Desai were all helpful. I thank Donghyun David Kim, Julie Kim and Seongjoon Kang who all helped me to analyze transgenic *Arabidopsis* plants.

Sunny and I were delighted to see our old friends, Dr. Yong Suk Cho, his wife Gumhee Baek and their lovely daughter, Yoonjeong, when we arrived here five years ago. Dr. Cho – the fresh Ph.D is proud to be called this way – is brilliant in genetics. I have benefitted from the infrequent discussions with him, which were triggered by a small quantity of fermented *Hordeum* sugar drink. Moreover, it was always fun to watch the overtly expressive Dr. Cho bring about instant, marvelous and mischievous events (or problems?) while the thoughtful Gumhee quietly keeps things straight. Priceless is every moment with them in many places around Austin, like zoos, parks, malls and bars. So I thank my good friends for keeping my family happy. They seem as if they are a part of my own family. In fact, Dr. Cho whishes my son to be his son-in-law someday. I will have to wait and see if his wish comes true.

It was comforting to have Dr. Yonghwan Kim and Eundeok Kim here, who came from my old lab in Seoul National University. Not only did they make my life in Austin a lot easier, but Yonghwan also shared his clever idea about DNA repair study and Eundeok helped to develop several experimental concepts, such as measuring starch in tobacco. Dr. Youngsam Lee gave a short, yet, informative lesson about X-ray crystallography when I discovered beautiful salt crystals in my aged protein preparation. I also thank them and their families for staying close to my family.

I was blessed to teach with great professors like Drs. John La Claire, Julie Palmer, James Mauseth, Robert Fulginiti, Tanya Paull, Suzanne Barth, Ruth Franks, and Inder Saxena. They showed me how to develop a course, interact with students, and deliver enthusiasm as well as knowledge. I am confident to teach my own students elsewhere in future since all the professors prepared me very well. Especially, Dr. La Claire with his superb course, both in lecture and labs, inspired me to become a better teacher.

I owe many thanks to the graduate coordinator, Tamra Rogers, for remaining so helpful since the beginning of my graduate school. She took good care of my interaction with the department and the school. The green house manager, Dena Sutton, and her staff were particularly good, of course, at keeping my plants from desiccation, malnutrition and pests. A lot of my plants would have been ruined without their invaluable help.

It was fortunate to know Drs. Dwight Romanovicz and Angela Bardo. They were very effective in getting my microscopy works done. The witty Dr. Romanovicz used to make fun of me by calling me a ‘gel’ biologist. If I make any big or small scientific

contribution to the society with a microscope, it is because he taught me very well and was successful in transforming me into a microscopist. Eventually, he started to acknowledge me as a true 'cell' biologist.

An erudite professor, Mona Mehdy, has always backed me up with her bottomless knowledge. Along with Dr. La Claire, I should have her in my doctoral committee. She also helped me troubleshoot *Arabidopsis* genomic DNA analysis together with her student, Chenggang Liu. I also thank my fellow students in her lab, Songhita Das and Tharindu Weeraratne, for a fun time in graduate school. I appreciate Ling Zhu, Nidhi Sharma, and Dr. Hui Shen in the Huq lab for teaching me how to take care of and transform *Arabidopsis*. Without their help, I could not have started the chloroplast DNA repair study.

I am happy and sad to thank my parents for the first time in a long while. It took a lot longer than I expected. Now, you may call your son a doctor as long as you like.

I ask my parents-in-law for their forgiveness, for I took from them, their only daughter far, far away and never sent her back. Nevertheless, they have continued to pray and bless me. I sincerely appreciate them. I am also sorry for keeping their grandchildren, Youngmeen and Zimin far away. By the way, those young rascals have been very fond of using 'Mr. Watering Can' every holiday and weekend. It was their youthful magic and enthusiastic contribution to science: turning plant cultivation into a fun job, which could have been my jejune task otherwise. Above all, I thank them just for being my son and daughter.

A great thinker, Isaac Newton said that if he had seen a little further it was by standing on the shoulders of Giants. I have achieved nothing compared to him, but I too was looking at the world on the shoulders of my own giants. I am so blessed to find them. They are Drs. Stanley J. Roux, Robert K. Jansen, Enamul Huq, Arturo De Lozanne, Aravind Somanchi, David E. Graham and David L. Herrin. These wise giants were courageous to take an unintelligent student five years ago and persevered to change him more insightful by using their attentive guidance. Any minute success of mine is thanks to all the people mentioned above.

Ineffable is how great Dr. Herrin has been for me; his constructive criticism, enthusiasm, wisdom, support and encouragement were simply amazing. I dare not start to describe his contributions for I would never be able to finish. I am truly grateful for the years with him. So I devote my best four words to him: “Thank you very much!”

.

Consequences, Repair, and Utilization of an Induced Double-Strand Break in the Chloroplast DNA of *Arabidopsis* and Tobacco

Publication No._____

Taegun Kwon, Ph.D.

The University of Texas at Austin, 2011

Supervisor: David L. Herrin

In mature chloroplasts, the DNA (cpDNA) is surrounded by a potentially genotoxic environment that would make the mitochondrial DNA milieu look like a "nadree" (picnic). And yet, the slower evolution of cpDNA compared to other cellular genomes suggests that this organelle must have efficient mechanisms for repairing DNA. Unfortunately, those mechanisms have been barely noted, much less studied. This dissertation describes a novel approach that was developed to study how chloroplasts of *Arabidopsis* repair the most severe form of DNA damage, a double-strand break (described in Chapter 2). The success with this approach also prompted the development of a new method for site-specific modification of tobacco cpDNA that is described in Chapter 3.

To study the consequences and repair of a break in the circular plastid genome, we developed an inducible system based on a *psbA*-intron endonuclease from *Chlamydomonas* (I-*Cre*II) that specifically cleaves the *psbA* gene of *Arabidopsis*. The protein was targeted to the chloroplast using the rbcS1 transit peptide, and activation of the nuclear gene was made dependent on an exogenous inducer (β-estradiol). In *Chlamydomonas*, I-*Cre*II cleavage at *psbA* was repaired, in the absence of the intron, by homologous recombination between repeated sequences (20-60 bp) that are abundant in that genome. By comparison, *Arabidopsis* cpDNA is very repeat-poor. Nonetheless, phenotypically strong and weak transgenic lines were obtained, and shown to correlate with *I-CreII* expression levels. Southern blot hybridizations indicated a substantial loss of *psbA*, but not cpDNA as a whole, in the strongly-expressing line. PCR analysis identified deletions nested around the I-*Cre*II cleavage site that were indicative of repair using microhomology (6-12 bp perfect repeats, or 10-16 bp with mismatches) or no homology. The results provide evidence of alternative repair pathways in the *Arabidopsis* chloroplast that resemble the nuclear microhomology-mediated and nonhomologous end-joining pathways, in terms of the homology requirement. Moreover, when taken together with the results from *Chlamydomonas*, plus other considerations, the data suggests that an evolutionary relationship may exist between the repeat structure of cpDNA and the organelle's ability to repair broken chromosomes.

Taking advantage of the inducible I-*Cre*II system, I developed a method to delete defined regions of cpDNA in tobacco, which was named DREEM (for *d*irect *re*peat and *e*ndonuclease *m*ediated). Chloroplast transformation was used to introduce an I-*Cre*II

cleavage site adjacent to an *aadA:gfp* marker and flanked by a direct repeat of 84 bp. When chloroplast-targeted I-*Cre*II was induced with β-estradiol during germination, complete loss of the *aadA:gfp* marker occurred by SSA-type repair involving the 84-bp direct repeat. I obtained additional evidence for DREEM effectiveness by deleting 3.5 kb of native cpDNA that included part of the large *ycf1* gene. DREEM can be used for other modifications besides gene deletions, partly because it is seamless and leaves no trace of introduced DNA. Since expression of the endonuclease is controlled by steroid application (and concentration), and the deleted cpDNA is probably destroyed during the SSA process, this inducible gene-ablation technique could enable the study of essential chloroplast genes *in vivo*.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

A.U.	arbitrary unit(s)
BIR	break-induced replication
bp	base pair(s)
cfu	colony forming unit(s)
cpDNA	chloroplast DNA
DSB	double-strand break(s)
DSBR	double-strand break repair
DREEM	direct repeat and endonuclease mediated
EDTA	ethylenediaminetetraacetic acid
GMO	genetically modified organism(s)
GFP	green fluorescent protein
HR	homologous recombination
hr	hour(s)
IPTG	isopropylthio-β-galactoside
IR	inverted repeat
kb	kilobase pair(s)
LSC	large single copy
LB	Luria-Bertani
min	minute(s)
MS	Murashige-Skoog
MMEJ	microhomology-mediated end joining
NHEJ	non-homologous end joining
nt	nucleotide(s)
ORF	open reading frame
PCR	polymerase chain reaction
PSI	photosystem I
PSII	photosystem II
RIII	rbcS:I-*Cre*II
RIIIG	rbcS:I-*Cre*II:GFP
RM	medium for tobacco root regeneration
RMOP	medium for tobacco shoot regeneration
SDS	sodium dodecyl sulfate
SDSA	synthesis-dependent strand annealing
sec	second(s)
SSA	single strand annealing
SSC	small single copy
UTR	untranslated region
WT	wild-type

Chapter 1. Introduction

1.1 Overview

The chloroplast is a semi-autonomous cell organelle that carries its own genome, which likely evolved from an endosymbiotic cyanobacterium (Archibald 2009). It offered to me a question that this dissertation addresses: How does the chloroplast maintain its genomic integrity while carrying out the potentially genotoxic reactions of photosynthesis, and under direct sunlight, which is also harmful to DNA (Sakai *et al*. 2004)? In the begining, I learned that Dr. Herrin had conceived the idea of using a homing endonuclease (I-*Cre*II), discovered by his lab (Kim *et al*. 2005), to study repair of broken chloroplast DNA in land plants, especially *Arabidopsis*. The reasons for studying DNA repair in *Arabidopsis* were because it has a native I-*Cre*II cleavage site (Kim *et al*. 2005), and because there are plentiful genetic resources, such as a fully-sequenced genome, many mutant seed stocks, and well-defined experimental protocols (Swarbreck *et al*. 2008).

After having some success in using I-*Cre*II in *Arabidopsis* (described in Chapter 2), I decided to apply the same principle to tobacco, which has a well-established chloroplast transformation protocol (Lutz *et al*. 2007; Maliga 2004; Svab *et al*. 1990), unlike *Arabidopsis* (Sikdar *et al*. 1998). I was already trying to use intron homing and I-*Cre*II to improve tobacco chloroplast transformation (Odom *et al*. 2001), which is difficult even in that system. After enduring many frustrating months, the effort produced a novel method for conditional chloroplast-gene deletion. To verify the method, which

we call DREEM, I conditionally deleted much (3 kb) of the second largest chloroplast gene (*ycf1*), which is one of the most elusive in terms of function.

1.2 Background

1.2.1 Chloroplast structure and function

Chloroplasts are eukaryotic cellular organelles that presumably evolved from a free-living cyanobacteria-like organism via an ancient endosymbiosis (Allsopp 1969; Archibald 2009; Cavalier-Smith 2002; Gray and Doolittle 1982; Moreira *et al*. 2000). They are surrounded by an envelope made up of two membranes, which control molecular exchange with the cytoplasm. These membranes may have evolved from the cellular membrane of the endosymbiont, after the loss of the eukaryotic food vacuole (Cavalier-Smith 1982, 2000). A mature chloroplast also contains well-differentiated and fully photosynthetic thylakoid membranes that have an immense surface to volume ratio, and form granal stacks and intergranal lamellae (Anderson 1999; Mustardy *et al*. 2008). The granal stack and stromal lamellae are enriched for Photosystems II and I (and their associated light-harvesting complexes), respectively (Andersson and Anderson 1980). Between the inner envelope and thylakoid membranes is the stroma, where carbon fixation takes place and where chloroplast DNA (cpDNA), ribosomes and starch granules are found (Bassham *et al*. 1950; Suss *et al*. 1993). In recent years, evidence for narrow tubular stromules that interconnect chloroplasts and possibly allow for molecular exchange has been published (Hanson and Sattarzadeh 2008; Kohler *et al*. 1997). The

relative simplicity of the ultrastructural picture, however, contrasts with the molecular complexity of chloroplasts, which likely contains 3,000-5,000 proteins and many key metabolic pathways (Martin and Herrmann 1998).

Functionally, chloroplasts are the sites of: (*i*) photosynthesis, which produces cellular energy (ATP) and reducing power (NADPH); (*ii*) anabolic pathways that produce starch (Stitt and Heldt 1981), fatty acids (Roughan *et al*. 1979), amino acids (Elias and Givan 1977), pigments, and nucleotides (Cunningham and Gantt 1998; Stasolla *et al*. 2003; Von Wettstein *et al*. 1995); (*iii*) sulfur and nitrogen assimilation (Hell 1997; Lam *et al*. 1996); and (*iv*) cell signaling that is involved in gravity perception (Stanga *et al*. 2009), stomatal attitude (Tominaga *et al*. 2001), and the defense response to pathogen infection (Caplan *et al*. 2008). Thus, mutations, either in the nuclear (Robertson *et al*. 1995) or chloroplast (Archer and Bonnett 1987) genome, that influence chloroplast development often show pleiotropic effects on plant growth and development.

1.2.2 Chloroplast-encoded genes

Over a billion years of evolution, most of the endosymbiont's genes were lost or transferred to the eukaryotic host genome (Martin and Herrmann 1998; Martin *et al*. 2002; Timmis and Scott 1983). Based on genomics, about 3,000-5,000 proteins are estimated to be in the chloroplast (Abdallah *et al*. 2000; Jan van Wijk 2000; Martin and Herrmann 1998), but land plant chloroplasts carry only 100-130 genes in their DNA (Sugiura 1992). Most characterized chloroplast genes can be classified into two groups based on function (Sugiura 1992): First, about 40 genes encode proteins involved in

photosynthesis, such as subunits of Rubisco, cytochrome b_6/f complex, ATP synthetase, and the core reaction centers of photosystems I (e.g., *psaA* and *psaB*) and II (e.g., *psbA*); secondly, ~60-70 genes contribute to the genetic systems for transcription (core subunits of RNA polymerase), and translation (ribosomal proteins, ribosomal RNAs and tRNAs) (Sugiura 1992). A few other genes have different functions, such as *accD* for fatty acid synthesis (Sasaki *et al*. 1993) and *ycf5*/*ccsA* for attachment of heme to *c*-type cytochrome (Xie and Merchant 1996), and there remain several conserved open reading frames (the *ycf* genes) of unknown functions (Rochaix 1997; Bock 2007). Among them are the two very large (~2,000 amino acids) genes, *ycf1* and *ycf2*, which were resistant to knock-out experiments, implying they are indispensable for plant growth (Drescher *et al*. 2000).

1.2.3 Chloroplast DNA size and organization

The cpDNA of land plants is typically a 120-200-kb circular molecule (Kolodner and Tewari 1975) – although this was questioned recently (Bendich 2004) – that exists mainly as a monomer, but with several multimeric circles (Deng *et al*. 1989; Kolodner and Tewari 1979; Lilly *et al*. 2001). The typical chloroplast genome has two large inverted repeats that separate the small single-copy region from the large single-copy region (Goulding *et al*. 1996). Because of recombination between the inverted repeats, there are two cpDNA isoforms *in vivo* that differ in their relative orientation of the single-copy regions, but are identical otherwise (Palmer 1983). Chloroplasts are usually highly polyploid, with each organelle having 30-100 copies of the genome (Scott and Possingham 1980). Since there are ~100 chloroplasts in a mature mesophyll cell, there

can be thousands of copies of cpDNA, and it can constitute up to 20% of the total DNA in a leaf (Scott and Possingham 1980; Thomas and Rose 1983). The polyploidy of the chloroplast genome acts as a barrier to the propagation and fixation of mutations, and can partially account for its highly conservative evolution (Clegg *et al*. 1994). In practice, it also makes genetic manipulation of the organellar genome difficult.

CpDNA is organized into 5-10 nucleoids per organelle (Kuroiwa 1991) and lacks histones. However, several nucleoid proteins have been identified (Jeong *et al*. 2003; Kobayashi *et al*. 2002; Sato *et al*. 1993), and suggested to have roles in DNA compaction (Chi-Ham *et al*. 2002) or transcriptional repression (Sato *et al*. 1998), but their purported functions have not been proven genetically (Heinhorst *et al*. 2004). The intra-organellar location of nucleoids changes during plastid development from the envelope membrane to the thylakoid membrane in mature chloroplasts (Heinhorst *et al*. 2004; Sato *et al*. 1997)

1.2.4 General DNA damage and repair in chloroplast

Due to its close proximity to the site of oxygen evolution and electron transport (Liu and Rose 1992; Sato *et al*. 1993), cpDNA is more prone to damage from reactive oxygen species (singlet oxygen, superoxide, hydroxyl radical, and nitric oxides) than the nuclear or mitochondrial genomes (Allen and Raven 1996). High-energy radiation in the full spectrum of sunlight is also a major source of DNA damage (Beukers and Berends 1960; Scholes *et al*. 1960). Several mechanisms are known to exist in chloroplasts to cope with the diversity of DNA damage (Britt 1996). Light-dependent DNA repair by a

photolyase resolves cyclobutane pyrimidine dimers that are caused by UV irradiation (Cannon *et al*. 1995; Selby and Sancar 2006). There is also biochemical evidence for base-excision repair enzymes in the corn (Bensen and Warner 1987) and *Arabidoposis* (Gutman and Niyogi 2009) chloroplasts that could potentially correct base damage (e.g., deamination, alkylation and/or oxidation); however, gene knock-outs had no phenotype, raising concern over the significance of these enzymes (Gutman and Niyogi 2009). Homologous recombination occurs in chloroplasts (Cao *et al*. 1997; Cerutti and Jagendorf 1993; Khakhlova and Bock 2006), and seems to play a major role in DNA repair and replication (Cerutti *et al*. 1992; Kmiec *et al*. 2001; Odom *et al*. 2008), but the mechanisms are not well understood. Evidence for non-homologous recombination that, for example, might be used to repair breaks in cpDNA is basically lacking.

1.2.5 Repair of double-strand breaks in DNA

The most severe form of DNA damage is a double-strand break (DSB), which can result in gene loss, blocked replication or even apoptosis. DSBs are caused by ionizing radiation, oxidation by reactive oxygen species (Halliwell and Aruoma 1991), single-strand breaks during DNA replication (Cromie *et al*. 2001; Wyman and Kanaar 2006), genotoxic chemicals such as bleomycin (Claussen and Long 1999), and enzymes (topoisomerase and endonucleases). Repair mechanisms for DSBs are usually classified into two basic types: homologous recombination (HR) and non-homologous end-joining (NHEJ).

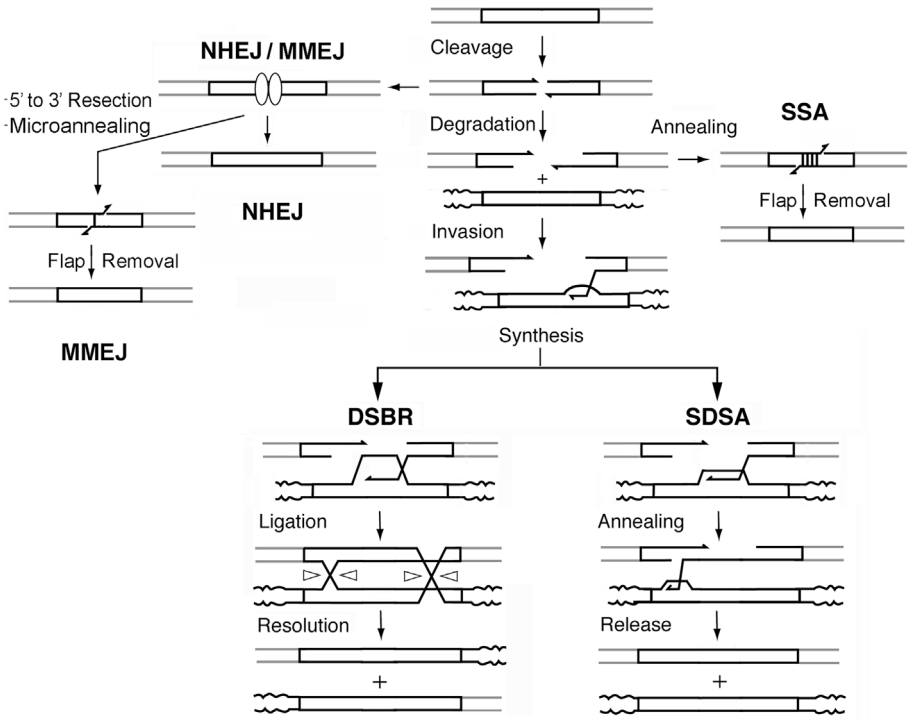


Figure 1. Pathways for repairing a double-strand break in DNA.

The SSA, DSBR, and SDSA pathways are considered to be a form of homologous recombination, because they require considerable regions of homology between the partner DNAs to be effective. The NHEJ pathway requires end-binding proteins (indicated with ovals), whereas MMEJ is an alternative or backup pathway that does not depend on those proteins. Because MMEJ is mediated by regions of micro-homology in the substrate DNAs, and involves the resection of DNA ends, it shares some features with the SSA pathway, yet it is considered to be a distinct pathway. Adapted from Figure 1 in Odom *et al* (2008).

1.2.5.1 Repair of DSBs by homologous recombination

Homologous recombination (HR) uses substantial homology to repair the broken DNA, and HR is dominant over NHEJ in yeast (Mezard and Nicolas 1994). Three distinctive pathways are known: double-strand break repair (DSBR), synthesis-dependent strand annealing (SDSA), and single-strand annealing (SSA) (Figure 1). The three

pathways start with resection at the 5'-ends of the broken DNA, resulting in single-stranded 3'-tails. The 3'-tails are recombinogenic and can invade an intact, homologous DNA duplex in the same, or in a homologous, chromosome. When that happens, repair can proceed by one of two possible repair pathways, either DSBR or SDSA.

Prior to that decision, however, the MRX (Mre11-Rad50-Xrs2) complex in yeast (Chen *et al*. 2001), or the MRN (Mre11-Rad50-NBS1) complex in mammals (Hopfner *et al*. 2002), senses the broken ends and tethers them together. Mre11 may help process the broken ends via its 3'-to-5' exonuclease activity, but it does not catalyze 5'-end resection (Llorente and Symington 2004); Exo1 catalyzes resection of 5'-ends in mammals (Nimonkar *et al*. 2008; Tsubouchi and Ogawa 2000). Replication protein A (RPA) binds the single-stranded region of the 3'-tails, and recruits Rad51, the main recombinase. Rad51 displaces RPA, forming a Rad51-nucleofilament (Song and Sung 2000). Rad51 is a homologue of bacterial RecA (Basile *et al*. 1992), and it catalyzes the homology search and invasion of the intact duplex by 3' single-stranded DNA (Sung 1994).

In the classic DSBR pathway (Szostak *et al*. 1983), the Rad51-nucleofilament invades the target duplex, displacing the homologous strands. Both strands of the intact DNA serve as templates for DNA synthesis, primed by the 3'-ends of the broken DNAs. (see Figure 1). At completion of DNA synthesis and ligation, two Holliday junctions are formed, and equal numbers of crossover and non-crossover products are produced depending on their resolution. The resolvase that cleaves the 4-stranded Holliday junctions has not been identified in most eukaryotes. However, the Rad51 paralogues, XRCC3 and Rad51C, are known to be involved in resolution of the junctions (Liu *et al*.

2004). An important point is that with DSBR (and SDSA), gene conversion occurs when the exchanged sequences are not completely homologous (the broken DNA becomes the converted sequence).

In the SDSA pathway, a single 3'-tail of the broken DNA invades the homologous target, forming a D-loop as it is copied (Figure 1). As DNA synthesis passes the break site, the nascent strand is released, and it hybridizes with the complementary strand of the break-partner. Holliday junctions are not formed, so crossovers do not occur, but gene conversion may take place if the target and invading sequences are not identical. This pathway explains the observation that non-crossover products are often more frequent than crossover products (Allers and Lichten 2001). Recently, it has been shown that the Srs2 helicase may be involved in the unwinding and re-annealing of the elongated strands during SDSA (Blanck *et al*. 2009; Ira *et al*. 2003; Robert *et al*. 2006).

In the SSA pathway, which is classically intramolecular (Lin *et al*. 1984), a 5'-exonuclease degrades the two broken DNA ends, exposing single-stranded 3'-tails (Figure 1). The two single strands can hybridize when homologous sequences (i.e., direct repeats) of sufficient length are exposed. Then, a flap endonuclease removes the protruding non-complementary tails, and a DNA ligase seals the nicks (Figure 1). The result is that DNA between the homologous repeats, and one of the repeat pairs are deleted. The SSA pathway can have great significance in DSB repair, especially in repeat-rich genomes like the *Chlamydomonas* cpDNA (Odom *et al*. 2008). In yeast, Rad52 and Rad59 are needed to promote the annealing step (Sugawara *et al*. 2000), and the Rad1/Rad10 endonuclease removes the 3' flaps (Davies *et al*. 1995; Fishman-Lobell

and Haber 1992). Also in yeast, the Msh2, Msh3 and Slx4 genes play roles in regulation of the SSA pathway (Flott *et al*. 2007; Saparbaev *et al*. 1996).

1.2.5.2 Repair of DSBs by non-homologous end-joining

Non-homologous end joining (NHEJ) directly ligates the two broken ends using little, or in some cases, no homology. It is the dominant DSB repair pathway in plant (Siebert and Puchta 2002) and mammalian nuclei (Sargent *et al*. 1997). NHEJ uses a sophisticated machinery to protect the broken ends, process them, if necessary, and re-ligate them together. Based on the early studies, which used DNA substrates with incompatible ends, NHEJ was considered to be highly error-prone, because the repaired DNA often had a small insertion or deletion. However, when the fidelity of the NHEJ pathway was evaluated using compatible ends, the error rate was estimated to be less than 0.1% of all end-joining events (Moore and Haber 1996; Pardo *et al*. 2009; Wilson *et al*. 1997).

For NHEJ, the Ku70/Ku80 heterodimer is a key player; it binds the free ends and protects them from degradation (Chen *et al*. 2001; Lee *et al*. 1998; Zhang *et al*. 2007). The MRX complex in yeast (mentioned above) also helps tether the broken ends together (Trujillo *et al*. 2003), although that role is not evident for its mammalian counterpart, MRN (Pardo *et al*. 2009). Instead, in mammals, a DNA-dependent protein kinase (DNA-PK) bridges the DSB, and regulates transcription factors like HOXC4, Oct-1 and Oct-2 (Gottlieb and Jackson 1993; Schild-Poulter *et al*. 2001). Incompatible ends are polished by other proteins, as needed: e.g., a polynucleotide kinase dephosphorylates 3' ends

(Chappell *et al*. 2002), nucleases resect some broken ends, Rad27 cleaves 5'-flaps (Tseng and Tomkinson 2004), and Artemis can remove both 5'- and 3'-flaps (Niewolik *et al*. 2006). The break is re-sealed by DNA ligase IV in mammals Dnl4 in yeast (Wilson *et al*. 1997). Ligation is known to be promoted by XRCC4 (Grawunder *et al*. 1997; Sibanda *et al*. 2001) and XLF (Ahnesorg *et al*. 2006) in mammals; Lif1 (Herrmann *et al*. 1998) and Nej1 (Kegel *et al*. 2001) are the respective yeast homologues.

Recently, microhomology-mediated end joining (MMEJ) has been recognized as a separate DSB repair pathway (Bennardo *et al*. 2008). Unlike NHEJ, MMEJ is independent of the Ku70/Ku80 proteins. However, it does rely on microhomology, i.e., direct repeats of 2-16 bp, which are too short for the SSA pathway (Ma *et al*. 2003). Similar to SSA, the DNA between the microrepeats gets deleted. In yeast, some proteins that mediate HR, such as Mre11-Rad50 and Rad1-Rad10 (endonuclease) are also required for MMEJ, although Rad51 and Rad52 are not (Decottignies 2007; Ma *et al*. 2003). In mammals, repair by MMEJ requires DNA ligases I and III, but not IV, which is the DNA ligase for NHEJ (Liang *et al*. 2008). In the nucleus of *Arabidopsis*, Mre11 also plays a role in MMEJ (Heacock *et al*. 2004).

1.2.5.3 Break-induced replication (BIR)

The aforementioned pathways repair DSBs that generate two free ends, whereas break-induced replication (BIR) repairs DSBs that have only one free end (Asai *et al*. 1994; Cromie *et al*. 2001; Wyman and Kanaar 2006). One-ended breaks are formed when a DNA replication fork encounters a nick on a template strand. In BIR, the 5'-strand of

the free end is degraded, and the resulting single-stranded 3'-tail invades a homologous DNA duplex. This turns into a DNA replication fork that copies the invaded DNA. Unlike the SDSA pathway above, DNA synthesis in BIR continues to the end of the chromosome, which can result in a massively long gene conversion tract (Cromie *et al*. 2001; Wyman and Kanaar 2006). In yeast, Rad51 catalyzes the homology search, and Rad52, Rad54, Rad55, and Rad57 are required (Davis and Symington 2004). There is also a Rad51-independent pathway – Rad52 and Rad59 promote strand invasion in the absence of Rad51 (Bai and Symington 1996).

1.2.5.4 DSB repair in bacteria

Only small subsets of bacteria are known to have Ku homologues, and are thus able to use the NHEJ-like pathway to repair a double-strand break (Bowater and Doherty 2006; Doherty *et al*. 2001). In *Mycobacterium*, Ku promotes joining of the two broken ends (Weller *et al*. 2002), and DNA ligase D processes and seals the break using its multiple enzymatic activities, which are nucleotidyl transferase, gap-filling polymerase, terminal transferase, primase, and 3' to 5' exonuclease (Della 2004). Knocking out homologous recombination in this organism showed that NHEJ mainly plays a role in DSB repair in late stationary phase (Stephanou 2007).

Recombinational repair of DNA in *Escherichia coli* is performed by the RecBCD and RecFOR pathways. RecBCD is more important in *E. coli*, but RecFOR has a broader phylogenetic distribution (Eggleston and West 1997; Eisen and Hanawalt 1999), and is the key to DSB repair in radiation-resistant *Deinococccus* (Bentchikou *et al*. 2010; Cox *et*

al. 2010; Rocha *et al*. 2005). RecBCD is a helicase-nuclease complex that resects the 5'-end of the broken DNA and loads RecA onto the single-stranded 3'-tail (Anderson *et al*. 1999; Anderson and Kowalczykowski 1997; Arnold and Kowalczykowski 2000). The *cis*-element known as Chi (for *c*rossover *h*ot *i*nstigator) directs a conformational change in RecBCD that switches it from degradative mode into recombinational mode (Anderson and Kowalczykowski 1997; Dixon and Kowalczykowski 1993).

In the absence of RecBCD, *E. coli* uses the RecQ helicase and a 5'-3' exonuclease (RecJ) to form the 3' tails (Ivancic-Bace *et al*. 2005). which will be bound by the single-strand binding protein, SSB. The concerted action of RecF, RecO, and RecR loads RecA onto the single-stranded tails (Ivancic-Bace *et al*. 2003), which then catalyzes the invasion of homologous double-stranded DNA. RecO is functionally related to Rad52, in that it promotes annealing of SSB-coated DNA (Umezu and Kolodner 1994). RecN, like RecR, is an ATP-binding protein that promotes complex-assembly by tethering the 3' ends together (Sanchez and Alonso 2005), analogous to Rad50 (Rocha *et al*. 2005; Wyman and Kanaar 2006). RecR from *Deinococcus* has a ring-like structure, and may clamp RecF onto DNA at the double strand-single strand junctions, in order to limit RecA filaments to the single-strand region (Honda *et al*. 2008). Also in *Deinococcus* the UvrD helicase is more important than the RecQ helicase in promoting repair by SDSA (Bentchikou *et al*. 2010; Cox *et al*. 2010). Finally, RuvA and RuvB (helicase) function together to promote branch migration (Iwasaki *et al*. 1992), as does the helicase RecG, which also has resolvase activity (Lloyd 1991; Lloyd and Sharples 1993). However,

RuvC is the endonuclease that resolves Holliday junctions formed by the classical DSBR pathway in *E. coli* (Shah *et al*. 1994).

1.2.5.5 DSB repair in the chloroplast

Chloroplast DSB repair is known to be involved in the homing of Group I introns in *Chlamydomonas* (Odom *et al*. 2001). Homing occurs when an intron inserts itself into an intron-less allele. An intron-encoded endonuclease makes a DSB in the intron-less target, and the intron gets copied into the intron-less allele during the repair (Belfort *et al*. 2002). Using a homing assay for the *psbA4* intron, evidence for all 3 major pathways of HR repair was obtained for the *Chlamydomonas* chloroplast (Odom *et al*. 2008). When the authors used the mobile intron with homology on both sides of the break as a probe, it got copied into the break by both the DSBR and SDSA pathways. In the absence of the invading intron, the DSB was mainly repaired by the SSA pathway using direct repeats of 20-60 bp that are abundant in that genome (Maul *et al*. 2002; Odom *et al*. 2008).

Compared to *Chlamydomonas*, the cpDNA of land plants has much less redundancy. Also, most of the direct repeats that are present are very small (< 12 bp) (Maul *et al*. 2002). This raised the question of how, or even if, land plant chloroplasts would repair a DSB in the genome.

1.3 Research objectives of this dissertation

The first objective was to investigate how a specific DSB in *Arabidopsis* cpDNA is repaired. The DSB was generated by expressing a plastid-targeted I-*Cre*II, which cuts a

native cleavage site in the *Arabidopsis psbA* gene. The homing endonuclease was made chemically inducible, and *Arabidopsis* cpDNA was analyzed by Southern hybridization and PCR, followed by DNA sequencing. The data provide clear evidence for repair of this repeat-deficient genome, and with some surprising results in terms of the mechanisms involved (Chapter 2).

The second objective was to develop a new method to modify cpDNA, based on the DSB repair pathways that had been studied in the lab. Tobacco (*Nicotiana tabacum*) was used for this research objective, because of its well-established chloroplast transformation protocol. An I-*Cre*II cleavage site was introduced in tobacco cpDNA next to a marker gene that was flanked by a direct repeat. Controlled I-*Cre*II expression from the nucleus was used to cut the target in cpDNA, which triggered a seamless deletion of the introduced marker gene (Chapter 3). I also deleted ~3.5 kb of the *ycf1* gene (tobacco *ycf1* is 5.7 kb) to further demonstrate the technique, which we call DREEM (for *d*irect *re*peat and *e*ndonuclease *m*ediated).

Chapter 2. Repair of an induced double-strand break in the chloroplast genome of *Arabidopsis*[*]

2.1 Introduction

Chloroplast DNA (cpDNA) is closely associated with the photosynthetic membranes that harvest radiant energy and strip electrons from water, producing molecular oxygen as a by-product (Rose 1979; Sakai *et al*. 2004). Highly reactive forms of oxygen are also generated, and despite the presence of detoxifying enzymes, photooxidative damage is a serious problem. Although there has been extensive study of the damage, protection, and repair of photosynthetic membranes, there has been little attention paid to the genetic consequences of photooxidative stress (GuhaMajumdar and Sears 2005). This can be attributed, at least in part, to a grossly incomplete knowledge of how chloroplast DNA is replicated and maintained throughout the life of a plant (GuhaMajumdar and Sears 2005; Maréchal *et al*. 2009).

However, despite the high-stress environment, the evolution of cpDNA is more conservative than nuclear or mitochondrial DNA (Korpelainen 2004). We can thus infer that DNA repair mechanisms play important roles in protecting chloroplast DNA, and are intimately involved in its evolution. And though some processes have been identified (Dürrenberger *et al*. 1996; Gutman and Niyogi 2009; Khakhlova and Bock 2006; Kmiec

[*] Large portions of this chapter have been previously published as: Microhomology-mediated and nonhomologous repair of a double-strand break in the chloroplast genome of *Arabidopsis*, by Taegun Kwon, Enamul Huq, and David L. Herrin, (2010) *Proceedings of the National Academy of Sciences USA* 107(31): 13954-13959.

et al. 2001; Selby and Sancar 2006; Small and Greimann 1977), the panoply of organelle repair mechanisms is not known, nor is it clear which are most important, or how they have coevolved with the genome. While it can be expected that some cpDNA repair processes will reflect a prokaryotic ancestry, the consequences of eukaryotic cell integration and over one billion years of evolution have surely produced some unique modifications to the repair machinery.

The most severe form of DNA damage is a double-strand break (DSB), which can result in gene loss, stalled DNA replication, and cell death (Paques and Haber 1999). DSBs are caused by ionizing radiation, chemicals, oxidation, enzymes, and single-strand breaks during replication (Paques and Haber 1999; Ragu *et al*. 2007). Mechanisms of DSB repair have been divided into two basic types, nonhomologous end-joining (NHEJ) and homologous recombination (HR) (Figure 1). NHEJ is the dominant nuclear response in animals and plants – it does not require homologous sequences, but is often mutagenic (Bleuyard *et al*. 2006; Wyman and Kanaar 2006). Repair by HR requires substantial homology, but the two pathways that use an intact chromosome to repair the broken one, DSBR and SDSA, are highly accurate. The third HR pathway, single-strand annealing (SSA), occurs between direct repeats longer than 30 bp and results in deletions; in plants, it is favored over the other two HR pathways (Puchta 2005).

In recent years, microhomology-mediated end joining (MMEJ) has been recognized as a distinct type of DSB repair in eukaryotes. Only very short (2-16 bp) regions of homology are needed for this pathway, and it typically leaves deletions like SSA. It has also been distinguished genetically from the HR pathways and classic NHEJ,

and in mammals it acts as a backup to NHEJ (Bennardo *et al*. 2008; Haber 2008; Heacock *et al*. 2004).

Group I intron homing involves the repair of a DSB induced by the intron's rare-cutting endonuclease, and results in gene conversion (Dürrenberger *et al*. 1996). Recently, our laboratory used a homing intron and its endonuclease (I-*Cre*II) to examine DSB repair in the *Chlamydomonas* chloroplast. Evidence of all three HR pathways was obtained, but not for NHEJ or MMEJ (Odom *et al*. 2008). In the absence of the intron, repair of the DSB was mostly by SSA between intergenic repeats of 20-60 bp that are abundant in that chloroplast genome (Maul *et al*. 2002). Compared to *Chlamydomonas*, however, the chloroplast DNA of land plants, especially *Arabidopsis*, is repeat-poor (Maul *et al*. 2002; Sato *et al*. 1999). For example, Reputer (Kurtz *et al*. 2001) analysis of the 155-kb *Arabidopsis* genome revealed only 31 direct repeats longer than 29 bp, and half of them map to the same small region of the genome (GenBank NC_000932). Thus, it was of interest to see how a land plant chloroplast would handle a similar DSB, while also recognizing the potential for further study in *Arabidopsis*. To this end, an inducible system was developed based on I-*Cre*II, which I targeted to the chloroplast with the transit peptide of *rbcS1*. Once inside the organelle, it cleaved the endogenous *psbA* gene (Kim *et al*. 2005; Krebbers *et al*. 1988; Lee *et al*. 2006; Zuo *et al*. 2001).

2.2 Results

*2.2.1 Recombinant I-*Cre*II constructs to induce a double strand break in chloroplast*

A homing endonuclease, I-*Cre*II, which originated from the *psbA4* intron of *Chlamydomonas reinhardtii* was used to make a DSB in *Arabidopsis* cpDNA within the *psbA* gene (Kim *et al*. 2005). I-*Cre*II was targeted to the chloroplast with the 59-amino acid long transit peptide of *rbcS1* (Krebbers *et al*. 1988; Lee *et al*. 2006), creating a translational fusion construct, *rbcS:I-CreII*, also referred to as *RIII*. In an attempt to track I-*Cre*II in the plant, a *gfp* gene was fused to the carboxyl-terminus, resulting in the DNA construct, *rbcS:I-CreII:gfp*, also referred to as *RIIIG*. Both the *RIII* (without *gfp*) and *RIIIG* (with *gfp*) constructs were used to transform *Arabidopsis*, but since the initial results with each variant were comparable, I focused on plants with *RIIIG* (*rbcS:I-CreII:gfp*).

2.2.2 In vivo *and* in vitro *assay for endonuclease activity of the I-*Cre*II fusion proteins*

Before producing the transgenic plants, I verified that the fusion constructs had endonuclease activity using *in vivo* and *in vitro* assays. The *in vivo* test was a plasmid exclusion assay carried out in *Escherichia coli*. The assay is based on the fact that a target-site plasmid cannot be maintained in the same *E. coli* cells as an active homing endonuclease plasmid, even if the plasmids have different selectable markers (Gruen *et al*. 2002). The bacterial cells were sequentially transformed with a substrate plasmid (that is kanamycin resistant) and then with an expression plasmid (that confers ampicillin

resistance). Cells producing an active I-*Cre*II fusion protein will not survive double-selection with both antibiotics. It should be noted that the plasmids used for this assay have identical origins of replication. While it is generally to be avoided (to use plasmids in the same compatibility group), the plasimids can be stably maintained so long as both antibiotics are always present in the medium. The effect of the plasmid incompatibility in this assay appears is to make the doubly-transformed cells even more sensitive to I-*Cre*II activity. As shown in Figure 2, the substrate plasmids pDrail (with one I-*Cre*II cleavage site) and pDrais (with five cleavage sites) gave a high number of double-transformants only with the control expression vector (pRSETA) and not with either recombinant expression plasmid (pRSET-RIIIG or pRSET-RIII). The results also indicate that the respective I-*Cre*II–fusion plasmids coexist with the control substrate-plasmid that lacked any I-*Cre*II cleavage sites (pRey), suggesting that the endonuclease is not finding a cleavage site in the main *E. coli* chromosome. This result suggests that both the RIIIG (rbcS:I-*Cre*II:GFP) and RIII (rbcS:I-*Cre*II) proteins have I-*Cre*II–like endonuclease activity in *E. coli*.

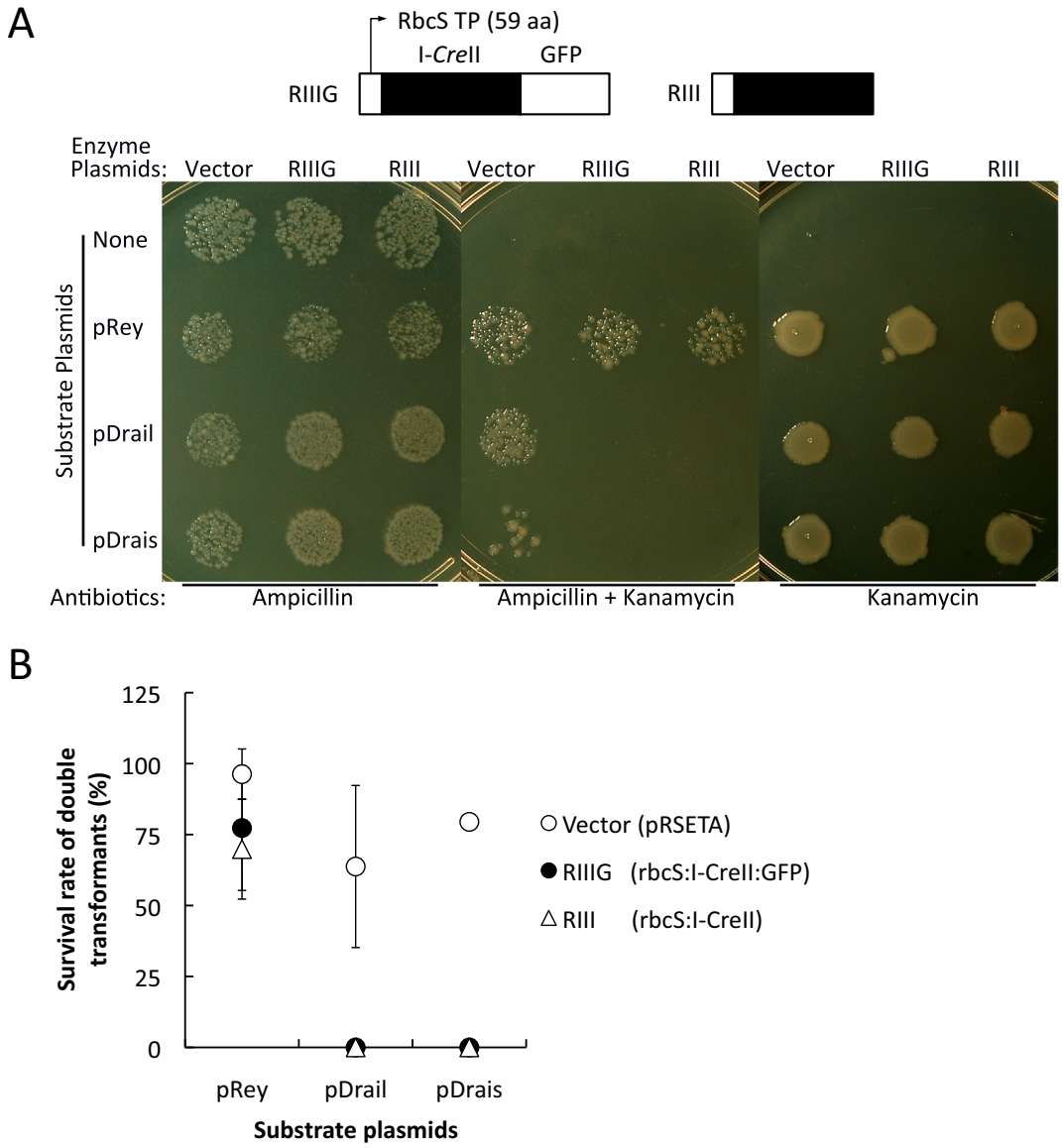


Figure 2. Plasmid exclusion assay for I-*Cre*II activity of the fusion proteins.

(**A**) *E. coli* strains carrying the substrate plasmid (*left*) (and which confers Kan^R), were transformed with the enzyme-expression plasmid (*top*) (which is Amp^R). Aliquots (5 μL) of each transformation were spotted on LB plates containing the indicated antibiotic(s) (*bottom*) and grown overnight. Schematic diagrams of the fusion proteins are on top of the photograph.

(**B**) Survival rate of the double transformants. Aliquots of the indicated transformations were diluted and plated on ampicillin and on both antibiotics (kanamycin and ampicillin). The survival rate was calculated as follow:

Survival rate (%) = (cfu·μg^{-1} DNA on kanamycin and ampicillin)/(cfu·μg^{-1} DNA on ampicillin) × 100.

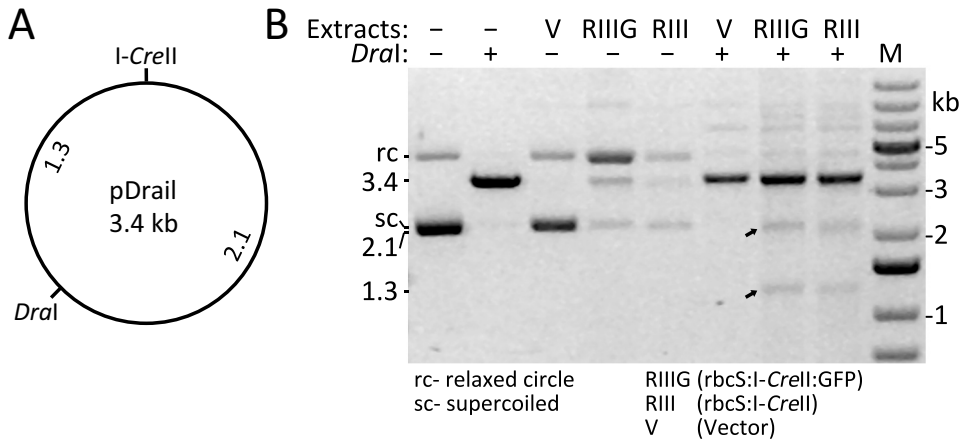


Figure 3. *In vitro* DNA cleavage with the I-*Cre*II fusion constructs.

(**A**) Simplified map of the substrate plasmid, pDrail, which has one I-*Cre*II site, and one *Dra*I site.

(**B**) Extracts (total soluble) were prepared from the *E. coli* strains (top) and added to the reactions with pDrail, as indicated. In some reactions, *Dra*I was added by itself or together with the indicated extract. Lane M contained DNA size markers. Other labels are explained below the agarose gel, whose image was digitally inverted. The small arrows (→) point to DNA bands that are expected from a double digestion of pDrail with *Dra*I and I-*Cre*II.

I-*Cre*II activity of the fusion constructs was further confirmed using bacterial extracts to cleave plasmid DNA *in vitro* (Figure 3). Crude extracts were prepared from the *E. coli* strains carrying the expression plasmids (pRSET-RIIIG, pRSET-RIII, and the parental Vector), and incubated with the substrate plasmid, pDrail. Only the extracts from the *E. coli* strains expressing the I-*Cre*II fusion genes contained specific cleavage activity toward the substrate (Figure 3).

Together with the plasmid exclusion assay, these data indicate that the homing endonuclease activity of I-*Cre*II was retained in the fusions with rbcS1 transit peptide and with or without GFP.

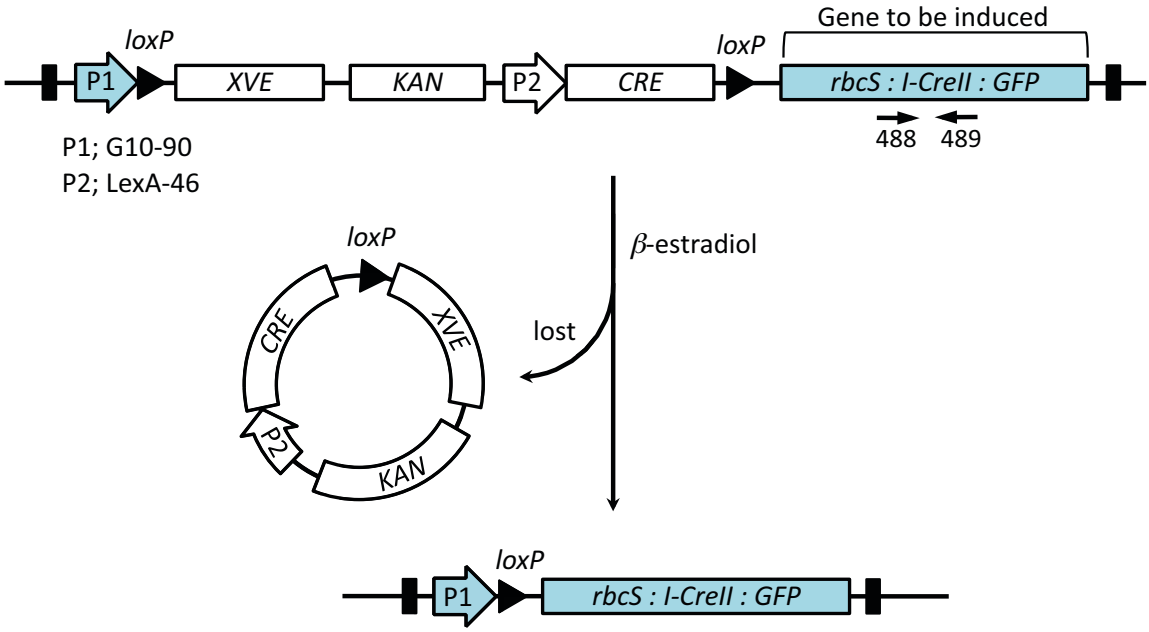


Figure 4. The inducible I-*Cre*II expression system.

In pX6RIIIG, the promoter-less *rbcS:I-CreII:gfp (RIIIG)* gene is separated from the constitutive promoter, P1, by the intervening sequence between two *loxP* sites, and is thus silent without induction. The chemical, β-estradiol, binds to and activates the artificial transcription factor XVE, which triggers expression of Cre recombinase. Cre-*loxP* recombination results in deletion of the DNA between two *loxP* sites, and expression of *rbcS:I-CreII:gfp*. The PCR primers (488 and 489) used for RT-PCR (Figure 5) are also indicated. The solid black rectangles are the T-DNA borders.

*2.2.3 Transformation of plants for an inducible, chloroplast-targeted I-*CreII

The I-*Cre*II endonuclease was expressed from the nucleus using chemical induction (Zuo *et al*. 2001); the application of β-estradiol induces transcription of the Cre recombinase (via the XVE receptor-activator), which catalyzes a reaction between loxP sites that places the G10-90 promoter upstream of the promoter-less target gene, *rbcS:I-CreII:gfp* (Figure 4). This vector was chosen because unscheduled expression of the endonuclease was expected to be lethal.

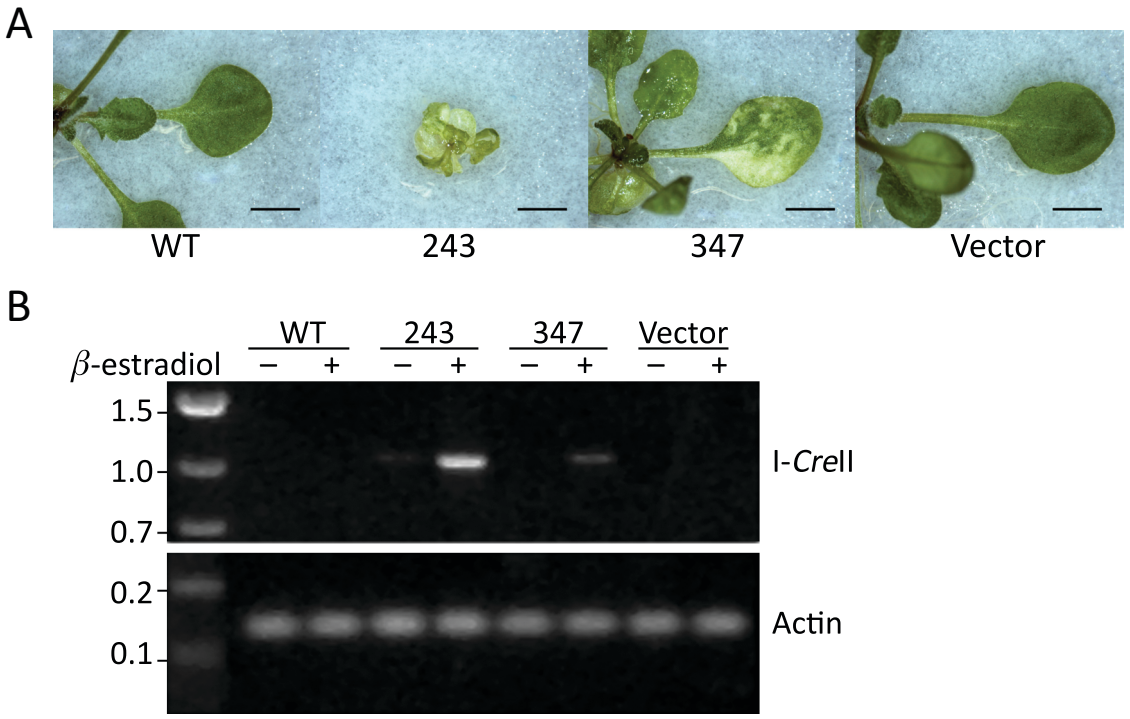


Figure 5. Transgenic *rbcS:I-CreII:gfp* plants germinated on induction medium for 2 weeks.

(**A**) The Vector line was transformed with the parental plasmid (pX6GFP) and lines 243 and 347 with pX6RIIIG (Figure 4). The seeds were germinated on induction medium (+ β-estradiol) for 2 weeks and representative photographs are shown. Scale bar = 3 mm.

(**B**) RT-PCR analysis of transgenic lines grown for 2 weeks on control (– β-estradiol) and induction (+ β-estradiol) medium. The RT-PCR primers were for the I-*Cre*II portion (488 and 489 in Figure 4) of *rbcS:I-CreII:gfp* (*top panel*), and an actin (Act2) control gene. Size markers are indicated to the left (kb). WT; wild type.

The *RIIIG* plasmid and parental vector were introduced into *Arabidopsis thaliana* (ecotype Col-0) using *Agrobacterium*-mediated transformation (Clough and Bent 1998). Healthy kanamycin-resistant lines possessing the T-DNA insertion were obtained, and subsequently germinated on β-estradiol (and control medium) to identify those that inducibly expressed *rbcS:I-CreII:gfp*. The inducible phenotypes could be classified as either strong or weak, and an example of each is shown in Figure 5A; compared to the

vector-only plants, the strong line (243) had very poor shoot development, whereas the weaker line (347) developed variegated leaves early, but then normally pigmented leaves thereafter. It should be said that the transgenic lines used in Figure 5A, and in the remaining analyses were homozygous for the T-DNA insertion, based on segregation analysis. The phenotypes, however, are similar to those observed in the original hemizygotes, because of the dominant nature of the transgene.

RT-PCR analysis indicated that both lines (243 and 347) displayed inducible expression of the I-*Cre*II-fusion transgene, but the signal was reproducibly greater, relative to the control (actin) mRNA, in the stronger 243 line (Figure 5B). No product was obtained with the vector line, as expected. The RT-PCR analysis also indicated that there was some basal expression of *rbcS:I-CreII:gfp* in line 243 without β-estradiol, but those plants did not exhibit a variegated phenotype like 347.

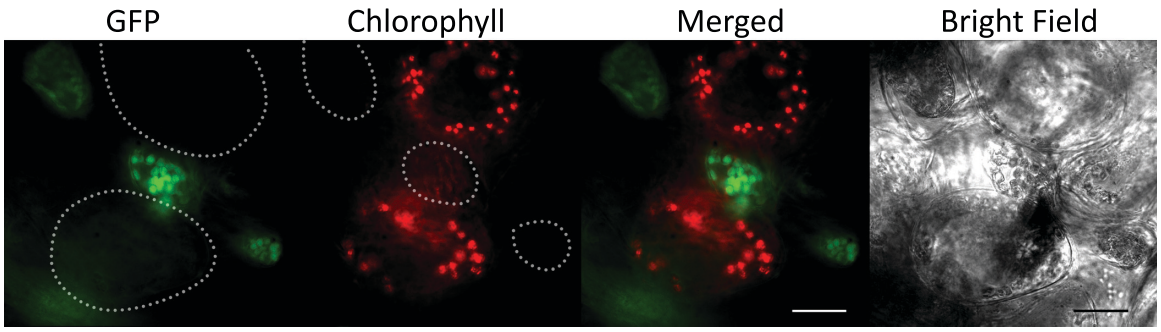


Figure 6. Fluorescence microscopy of a transgenic *rbcS:I-CreII:gfp* plant.

The images are of a cotyledon of a whole-mounted plant that was germinated on β-estradiol for 2 weeks. The dashed outlines on the GFP and chlorophyll fluorescence images are of cell boundaries. A merged image of the fluorescence images, and the corresponding bright field image are shown. Scale bar = 10 µm.

The level of GFP fluorescence was too low for microscopic detection in these induced lines (243 and 347), and in nearly all of the *rbcS:I-CreII:gfp* plants we examined, except for one that could not be sustained. The GFP fluorescence in that case was associated with plastids that lacked chlorophyll autofluorescence, indicative of severe damage (Figure 6). Most likely, plants that had enough *rbcS:I-CreII:gfp* expression to allow visualization of sustained GFP fluorescence were not viable.

*2.2.4 I-*Cre*II induced damage and repair at the* psbA *locus*

The recognition sequence for I-*Cre*II is an ~30-bp region in the highly conserved *psbA* gene (Kim *et al*. 2005), which is located in the large single-copy region, close to an inverted repeat (Sato *et al*. 1999) (Figure 7). Southern blot hybridizations of total DNA digested with restriction enzymes indicated only a slight loss of DNA at this locus in the 347 line, but a substantial (~70%) reduction in line 243, relative to nuclear 18S rDNA (Figure 7-8). Hybridizations with the *rpoA* gene, whose location in the chloroplast genome is far from *psbA* (Figure 7A), indicated the reduction in *psbA* signal was not due to a general loss of cpDNA as a whole, since the *rpoA*/18S rDNA ratio was similar in all 3 plant lines (Figure 7B. Thus, a preferential loss of *psbA* DNA had occurred in the stronger 243 line.

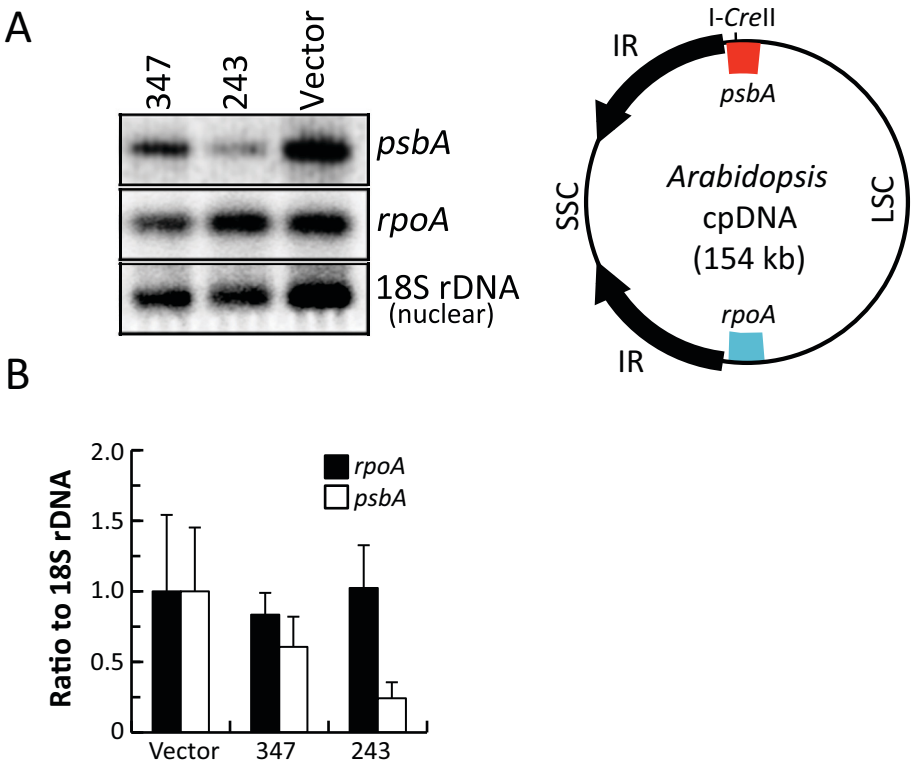


Figure 7. DNA hybridization analysis of transgenic plants expressing *rbcS:I-CreII:gfp*.

Equal amounts of DNA from the *rbcS:I-CreII:gfp* (347 and 243) and vector-transformed (Vector) lines, grown on β-estradiol, were restricted, Southern blotted, and hybridized to the indicated genes. The blots were hybridized to each probe sequentially, after stripping of the previous probe.

(**A**) Hybridization to *Sal*I digests. The bands were 9.6 (*psbA*), 11.4 (*rpoA*) and 6 kb (18S), respectively. The full-length blot images are available as Figure 8. The respective locations of the two plastid genes are indicated on the circular map. LSC, large single-copy region; SSC, small single-copy region; IR, inverted repeat.

(**B**) Quantification of the ratios of the *rpoA* and *psbA* signals to 18S rDNA. The ratios are means (± SD, *n* = 3) obtained by hybridization to *Sal*I, *Xho*I, and *Sal*I + *Xho*I digests.

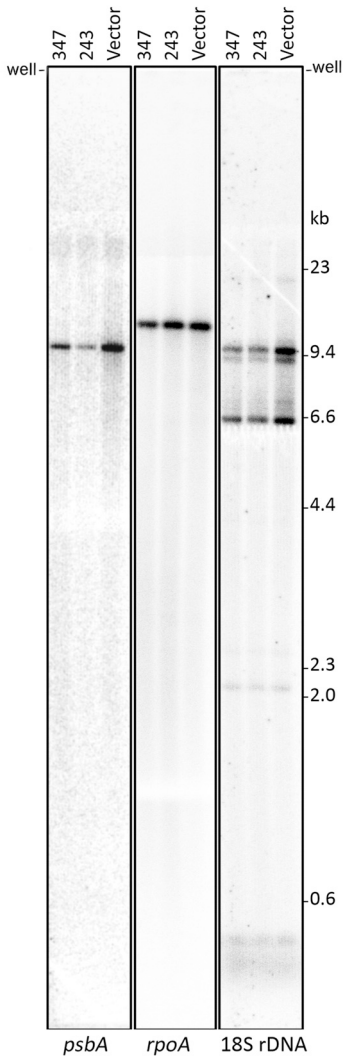


Figure 8. Southern blot analysis of transgenic plants expressing *rbcS:I-CreII:gfp*.

Equal amounts of genomic DNA isolated from the indicated (*Top*) plants grown on β-estradiol were restricted with *Sal*I and separated on a long (20 cm) agarose gel. After alkaline Southern blotting to a cationic nylon membrane, the blot was hybridized sequentially (after stripping off the previous probe) with gene-specific, ^{32}P-labeled DNA probes for *psbA*, *rpoA*, and nuclear 18S rDNA (described in Chapter 4). The blot probe is identified below its phosphorimage. The location of the bottom of the sample wells (the tops of the wells were removed, leaving the sides and bottom) is marked on both sides ("well") of the aligned images, and was verified from overexposed images. The positions and size (in kilobases) of ^{32}P-labeled size markers (*Hin*dIII-digested lambda phage) are indicated to the *right*.

To precisely characterize the deletions (or other possible changes) in line 243, trial PCR reactions were performed with nested primers that flanked the I-*Cre*II cleavage site and annealed ~5, 2.5, and 0.25 kb, respectively, from that site. Only the primers that annealed ~2.5 kb from the cleavage site were informative. They gave additional PCR products that were smaller than the 5-kb wild-type product, which was also correspondingly reduced in line 243 (Figure 9). The smaller products were not seen with the vector line. They were, however, faintly visible (on the original gels) with DNA from line 347, but to facilitate further analysis only PCR products from line 243 were sequenced.

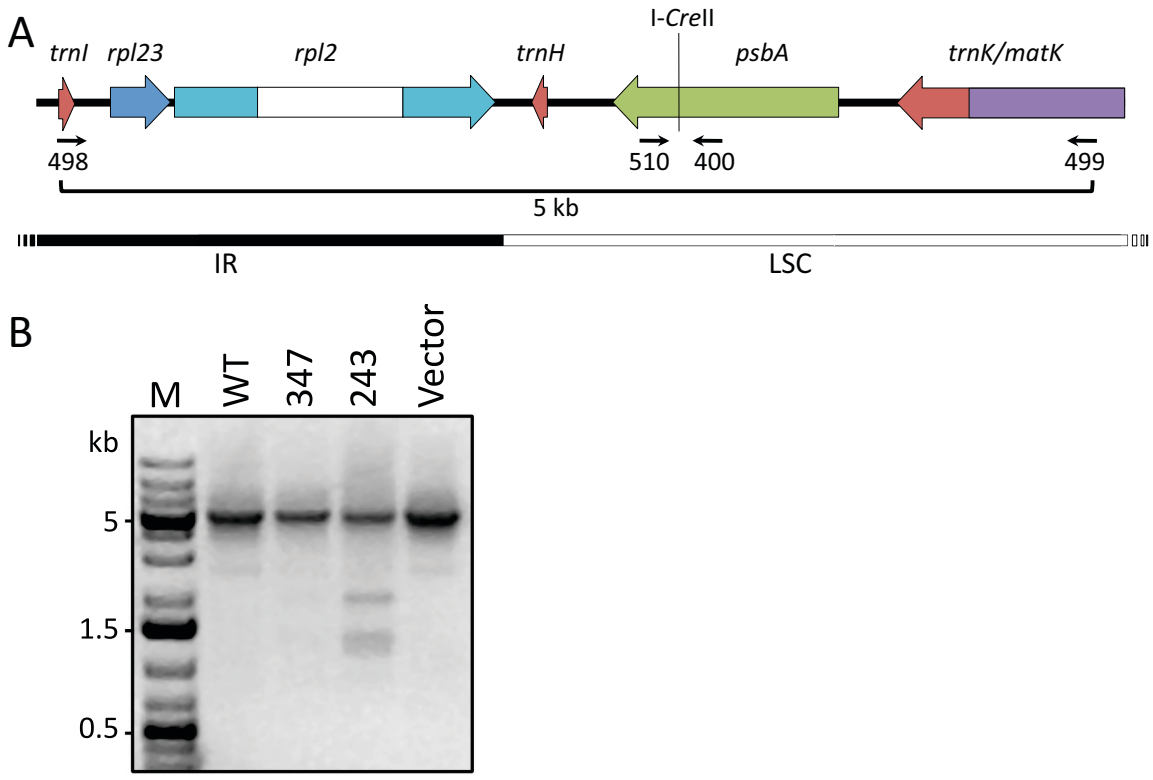


Figure 9. PCR analysis at the *psbA* locus of cpDNA in plants expressing *rbcS:I-CreII:gfp*.

(**A**) Map of the *psbA* region of cpDNA, with the PCR primers and I-*Cre*II cleavage site marked. MatK is the protein encoded within the *trnK* intron.

(**B**) Agarose gel of the PCR products. Total DNA from the plants grown on β-estradiol was used for PCR with the indicated primers (498 and 499). Lane M contained size markers, and the ethidium-DNA fluorescence image was inverted.

The major induced PCR products, which were 1.9, 1.3 and 1.2 kb (Figure 9B), were excised from gels for direct sequencing and cloning. Also, the total PCR products were gel-separated into 3 size fractions, which were cloned and sequenced. The results from all of this analysis are summarized in Table 1 and in Figure 10, which provides a map of the paired deletion-endpoints. First, it should be said that all of the mutant products had a deletion (no insertions were observed), which ranged in size from 3.1 to

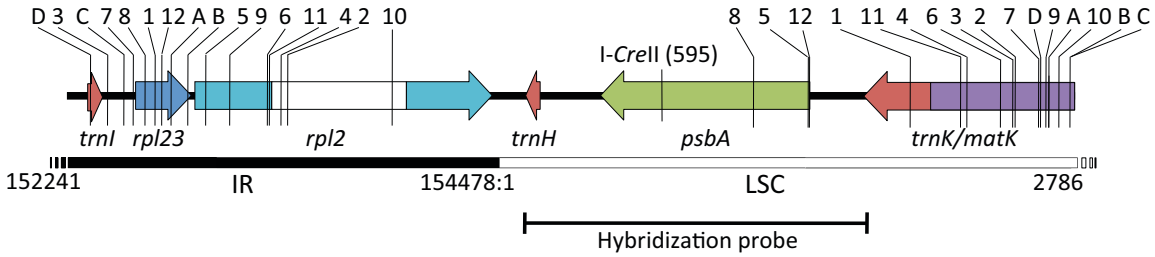


Figure 10. Map of the microrepeats and nonhomologous DNA junctions involved in repair of the DSB.

Microrepeats *1-12* and junctions *A-D* correspond to those in Table 1. The nucleotide numbering is from the *Arabidopsis* cpDNA (GenBank NC_000932), and the hybridization probe was used in Figures 7 and 8.

4.8 kb. The fact that the paired-endpoints for all these deletions map to either side of the I-*Cre*II site suggests strongly that they are associated with the DSB. It is also apparent from the map (Figure 10) that all the deletions encompassed the *trnH* gene and included at least 50% of the *psbA* and *rpl2* genes – in fact, *psbA* was completely deleted in most of them. These mutations, if accumulated to a sufficient level, would impact both plastid translation (*trnH* and *rpl2*) and photosynthesis (*psbA*), and could explain the poor shoot development in 243 (Ahlert *et al*. 2003).

It should be noted that the DNA from line 243 was also examined with primers that annealed close to the cleavage site (400 and 510 in Figure 9). The sequencing pattern of the PCR product was indistinguishable from wild-type, making it unlikely that the large 5-kb band (in Figure 9) contained many repair products with very short deletions or insertions.

The repair-junction sequences (Table 1) suggest that more than one process may be at work. Most of the repair events involved microhomology between the partners in

the form of direct microrepeats – the perfect repeats were only 6-12 bp (microrepeats *1-5*), whereas the imperfect repeats were 10-16 bp (microrepeats *6-12*); the longer imperfect repeats (15-16 bp) had two mismatches, whereas the shorter ones had only one. It is noteworthy that microrepeats 1, 6, and 8 correspond to the dominant sequences obtained from the three major PCR products (lane 243 in Figure 9). The fact that microrepeat *2* is highly represented in the sequenced clones (Table 1) probably reflects bias in the cloning process. The similarity of these repair junctions with those of the MMEJ pathway in the nucleus is striking (Bennardo *et al*. 2008; Glover *et al*. 2008; Haber 2008).

Several repair events were detected that could not be explained by microhomology between the DNA partners (junctions A-D in Table 1), and hence may represent a form of nonhomologous repair, like NHEJ. This seems to be a quantitatively minor pathway under these conditions. However, the large sizes of those deletions could indicate an important backup role to the microhomology-mediated pathway, which would be opposite to the situation in mammalian cells.

Table 1. Summary of microhomologous (microrepeats) and nonhomologous repair junctions.

Microrepeat*	PCR Product (kb)	Deletion (kb)	Repeat size / mismatches (bp)	Repeat position†	Sequence‡ IR-side repeat / Repair product(s) / LSC-side repeat	No. of clones
1§	1.2	3.8	12	152604	taggac**AGAAATAAAGCA**ttgggt	-
					taggac**AGAAATAAAGCA**cttttg	6
				1953	aataat**AGAAATAAAGCA**cttttg	-
2	1.3	3.7	11	153275	tcactg**ATCCAATTTGA**gtacct	-
					tcactg**ATCCAATTTGA**ttctaa	15
				2485	tttgct**ATCCAATTTGA**ttctaa	-
3	0.4	4.6	9	152364	aagtct**ATTGGAATT**ggctct	-
					aagtct**ATTGGAATT**ttgcta	1
				2471	gaagaa**ATTGGAATT**ttgcta	-
4	1.5	3.5	8	153244	aattag**GTTTACGA**cgaaac	-
					aattag**GTTTACGA**gccaaa	1
				2244	ttttgt**GTTTACGA**gccaaa	-
5	2.0	3.1	6	152862	cgtaga**CAGTCA**agtgaa	-
					cgtaga**CAGTCA**tggtaa	1
				1438	aaattg**CAGTCA**tggtaa	-
6§	1.3	3.7	16/2	153165	tataaa**AATGGGA**AA**TGCCCTA**cctttg	-
					tataaa**AATGGGA**AA**TGCCCTA**atacat	1
					tataaa**AATGGGA--TGCCCTA**atacat	3
				2410	tttact**AATGGGA--TGCCCTA**atacat	-
7	0.4	4.6	15/1	152493	tagggaAGAAAATCGATTTATggatgg	-
					tagggaAGAAAATCGATTTATattgac	0
					tagggaAGATAATTGATTTATattgac	2
				2602	atctttAGATAATTGATTTATattgac	-
8§	1.9	3.1	13/1	152553	tcggtt**ATTGGGG**A**AAAAAT**caatat	-
					tcggtt**ATTGGGG**A**AAAAAT**gcaatc	1
					tcggtt**ATTGGG**T**AAAAAT**gcaatc	6
				1158	tcccag**ATTGGG**T**AAAAAT**gcaatc	-
9	0.9	4.1	13/1	152982	gcgtct**ATACCG**T**AAAATA**gatttt	-
					gcgtct**ATACCG**T**AAAATA**acattg	0
					gcgtct**ATACCG**A**AAAATA**acattg	7
				2640	gaaacc**ATACCG**A**AAAATA**acattg	-
10	1.6	3.4	11/1	153807	gaaagc**CGTAT**G**CTTTG**gaagaa	-
					gaaagc**CGTAT**G**CTTTG**ttgcca	1
					gaaagc**CGTAT**C**CTTTG**ttgcca	0
				2704	aagcgg**CGTAT**C**CTTTG**ttgcca	-
11	1.5	3.5	11/1	153184	ggaaat**GCCC**T**ACCTTT**gagtgc	-
					ggaaat**GCCC**T**ACCTTT**ttaaaa	2
					ggaaat**GCCC**A**ACCTTT**ttaaaa	0
				2206	accaga**GCCC**A**ACCTTT**ttaaaa	-
12	1.7	4.4	10/1	152638	tttggt**GTCA**A**GGTAA**tagcta	-
					tttggt**GTCA**A**GGTAA**aatcct	1
					tttggt**GTCA**T**GGTAA**aatcct	0
				1440	attgca**GTCA**T**GGTAA**aatcct	-

Nonhomologous junction¶	PCR Product (kb)	Deletion (kb)	Position of junction nucleotide	IR-side Repair product LSC-side	No. of clones
A	0.6	4.4	152686	aaaggttaaa**A**gaatgggacc	-
				aaaggttaaa**AA**ttgccataa	1
			2655	cgaaaaataac**A**ttgccataa	-
B	0.6	4.5	152770	tccacctctt**A**gaaagaaaag	-
				tccacctctt**AA**tccttgagc	1
			2761	tgtatgaaagg**A**tccttgagc	-
C	0.2	4.8	152446	tcataacata**T**gaacagtaag	-
				tcataacata**TT**ccttgagca	1
			2762	gtatgaaagga**T**ccttgagca	-
D	0.2	4.8	152277	tccatggctg**A**atggttaaag	-
				tccatggctg**AT**tatattgac	1
			2614	agataattgat**T**tatattgac	-

* The repeats are the same as Figure 10.

† The numbers are from the chloroplast DNA sequences and refer to the first nucleotide of the repeat, which is in uppercase in the adjacent sequence column.

‡ IR and LSC refer to inverted repeat and large single copy region, respectively. Potential (and observed) repair products are between the repeats. In some cases (i.e. repeats with mismatches) only one of the two possible repair products were found. The repeat nucleotides and the nonhomologous junction nucleotides are in bold uppercase, and the mismatched nucleotides are in roman uppercase.

§ Microrepeats associated with the prominent PCR products in Figure 9.

¶ The junctions correspond to those in Figure 10.

2.3 Discussion

Transgenic homing endonucleases have been used extensively to study the repair of DSBs in nuclear chromosomes (Glover *et al*. 2008; Perez *et al*. 2005; Puchta 2005), but this is the first reported usage for a cell organelle. Transgenic restriction endonucleases have been employed to damage mitochondrial DNA in animal cells (Bacman *et al*. 2009), presumably because a homing enzyme target could not be introduced. Here, I used a relatively new homing endonuclease from a *Chlamydomonas* chloroplast intron (Kim *et al*. 2005), and retargeted it to the plastid. Since it cleaves the highly conserved *psbA* gene, I-*Cre*II could likely be used on many other plants. Finally, employing the steroid-inducible system (Zuo *et al*. 2001) may also have been important, especially considering the severe growth impairment of the higher-expressing line, 243.

Unlike intron homing, which alters the target DNA, repair of the DSB in this system by the same HR pathways would restore the target, which cannot be distinguished from DNA that was never cleaved. Hence, I can only infer repair events that have changed the target. Given the deficit of repeated sequences longer than 25 bp in *Arabidopsis* cpDNA, I was unsure how, or even if, the chromosome would become religated after the break. In the nucleus, repair with such limited or no homology occurs by pathways distinct from the HR pathways, and these processes were not known to occur in chloroplasts. However, these data clearly show that the organelle can repair a DSB using very limited homology, and with reasonable efficiency, based on the fact that there was little loss of cpDNA as a whole. Most of the repair events were mediated by

microhomology between the partner DNAs, reminiscent of nuclear MMEJ (Bennardo *et al*. 2008; Haber 2008; Heacock *et al*. 2004). Several repair events, however, lacked any mediating homology, indicative of an NHEJ-like mechanism (Gutman and Niyogi 2009; Wyman and Kanaar 2006). It is possible these latter DNAs are actually products of a microhomology-based machinery that can occasionally use non-homologous DNA ends. Alternatively, the nonhomologous products are from a distinct pathway that, based on deletion sizes, might back up the MMEJ-like pathway. Such a hierarchy would be the opposite of that in animal cells, where MMEJ backs up NHEJ (Bennardo *et al*. 2008).

An obvious question provoked by these experiments concerns the nature of the mediating proteins. The conserved Ku proteins, which are associated with NHEJ even in bacteria (Bowater and Doherty 2006), have not been found in chloroplasts, including *Arabidopsis* (Tamura *et al*. 2002; West *et al*. 2002). Moreover, the substantial size of these deletions and those in *Chlamydomonas* are consistent with the absence of strong end-protecting proteins like Ku. On the other hand, the deletions are suggestive of roles for a 5'-to-3' exonuclease, to resect the DNA and create 3' tails, and for an endonuclease to remove the 3' flaps (Haber 2008). Presumably, one or more proteins would be needed to promote microannealing, but it is not clear what any of these proteins are in the chloroplast. *Arabidopsis* should be a promising system in which to identify them, however.

Previously, the lack of evidence for NHEJ repair in *Chlamydomonas* (Odom *et al*. 2008), or associated with an insertion element in tobacco (Kohl and Bock 2009), prompted the suggestion that the absence of this ability might account for the lack of

horizontally acquired DNA in green plant chloroplasts. These results weaken that hypothesis, but stop short of negating it, since the NHEJ-like repair events were only a minor fraction of the total. It may also be relevant that we did not see any insertions of foreign (i.e., nuclear or mitochondrial) DNA in the break site; however, the laborious analysis of repair events limited the number that could be reasonably examined. Along this line, there are other factors that could limit horizontal transfer into cpDNA – e.g., the double-membrane envelope surrounding the chloroplast is a more formidable barrier than the nuclear envelope. Also, mitochondria have a peculiar tendency for fusion-subdivision cycles that might increase their exposure to foreign DNA (Logan 2010). Finally, there are examples of horizontally-acquired introns, and recently, whole genes in the cpDNA of unicellular green algae (Brouard *et al*. 2008; Odom *et al*. 2004). The equivalence of germ-line and vegetative cells in those organisms undoubtedly increases the probability of passing on horizontally-acquired DNA.

The commonly held notion that chloroplast DNA is circular rather than linear was challenged recently, based on physical analysis of DNA from isolated plastids (Bendich 2004). The stabilization and replication of linear DNA typically requires special structured end-sequences (e.g., telomeres), some of which are known from organelles (La Claire and Wang 2004; Vahrenholz *et al*. 1993). However, such sequences have not been identified for a chloroplast chromosome. These results thus support the notion that stabilizing end-structures would be needed to maintain linear chloroplast DNA.

Repetitive DNA is an important feature of all genomes, and its proportion can vary considerably (Shapiro and von Sternberg 2005). The cpDNA of *Chlamydomonas* is

relatively repeat-rich (Maul *et al*. 2002), and the I-*Cre*II-induced deletions described previously implicated perfect repeats of 15-62 bp, with a strong bias for repeats longer than 30 bp (Odom *et al*. 2008). Despite the paucity of such repeats in *Arabidopsis*, the I-*Cre*II-induced double-strand break was repaired efficiently, and the sizes of the accompanying deletions were similar to those obtained in *Chlamydomonas*. We suggested that evolution has endowed *Arabidopsis*, and probably other land plants, with the ability to repair DSBs without extensive homology, and that this ability may have impacted the evolution of cpDNA by reducing the need for sizable (> 20 bp) repeats (Kwon *et al*. 2010).

Chapter 3. Inducible deletion mutagenesis in tobacco chloroplasts using a new technique, DREEM (*d*irect *re*peat & *e*ndonuclease-*m*ediated)

3.1 Introduction

Intermolecular homologous recombination (HR) is the only reliable mechanism to achieve chloroplast transformation. Boynton and colleagues (1988) pioneered chloroplast transformation in *Chlamydomonas*, and tobacco (Svab *et al*. 1990) became the first land plant to be transformed by the same method. The list of transformed chloroplasts has been growing ever since: e.g., potato (Sidorov *et al*. 1999), tomato (Ruf *et al*. 2001), bladderpod (Skarjinskaia *et al*. 2003), soybean (Dufourmantel *et al*. 2004), carrot (Kumar et al. 2004a), cotton (Kumar et al. 2004b), petunia (Zubko *et al*. 2004), lettuce (Lelivelt *et al*. 2005), and cabbage (Liu *et al*. 2007). Typically, the transforming DNA contains the transgene(s) of interest and a selection marker, flanked by sequences homologous to the desired target site in cpDNA. Double HR events (crossovers) between the chloroplast-homologous sequences and cpDNA incorporate the transgene(s) into the chloroplast genome. The transforming DNA is delivered by particle bombardment of leaves, or freshly-grown, dispersed microcolonies of leaf cells (Huang *et al*. 2002), which are then subjected to multiple rounds of selection and regeneration. The sequential regeneration-selection cycles are necessary to achieve homoplastomic shoots and whole plants. *Nicotiana tabacum* (tobacco) is the only land plant whose chloroplast has been reproducibly transformed by a number of different groups [reviewed by Maliga (2004)],

and thus the plant I have used to develop a method for inducible deletion of chloroplast genes.

Chloroplast knock-out lines have been generated mainly by transforming the chloroplast with a plasmid that has the target gene containing the knock-out mutation (an insertion or deletion), flanked by long homologous wild-type sequences. A double-crossover in the flanking sequences replaces the wild-type gene with the knock-out (or mutant) gene, and after selection and regeneration, homoplastomic plants are obtained (Allison *et al*. 1996; Baena-Gonzalez *et al*. 2003; Kanevski and Maliga 1994; Kode *et al*. 2006; Legen *et al*. 2007; Tsuruya *et al*. 2006). If the target gene is essential, however, the regenerated shoots and plants remain heteroplastomic (with both wild-type and mutant copies of cpDNA), and often without a clear phenotype, such as the attempted knockouts of *ycf1* and *ycf2* (Drescher *et al*. 2000).

A variation of this approach is based on the chloroplast's ability to enact HR between direct repeats. Intramolecular recombination between neighboring direct repeats in cpDNA results in deletion of the DNA between the repeats. During tobacco chloroplast transformation (Staub and Maliga 1994), extrachromosomal multimeric minicircles (868 bp) were formed by homologous recombination between imperfect repeats (16 bp) in the intron sequences of *trnA* and *trnI*. The circular molecules were eventually lost after several months, but since only a fraction of the cpDNA molecules recombined in this way, there was no phenotypic affect on the plants.

A measure of recombination frequency in the *Chlamydomonas* chloroplast genome was obtained by using chloroplast transformation to duplicate part of a non-

essential gene that gives a visible phenotype. A portion (216 bp) of the *chlL* gene, which is required for chlorophyll synthesis in darkness, was added to the genome to form a direct repeat with part of the endogenous *chlL* gene. The *chlL* gene had also been disrupted with an *aadA* expression cassette, and so recombination between the direct repeats deleted *aadA* and restored *chlL*. This allowed a measurement of recombination frequency and implicated a role for a *recA*-like protein (Cerutti *et al*. 1995).

Repeat-mediated deletion has been used to make marker-free transplastomic plants. In the work of Iamtham and Day (2000), a histochemical marker (*uidA*), an antibiotic resistant marker (*aadA*), and a herbicide resistance marker (*bar*) were used for tobacco chloroplast transformation. Two of the markers had the identical *rrn16* promoter that formed a direct repeat of 174 bp, and all three markers had *psbA* 3'-UTR sequences that formed direct repeats of 418 bp. The construct was transformed into the *rbcL*/*accD* intergenic region. Intramolecular recombination events between the repeats yielded multiple outcomes, which included (*i*) deletion of *aadA*, (*ii*) deletion of *bar*, (*iii*) deletion of *uidA* and *aadA*, and/or (*iv*) deletion of *aadA* and *bar*, respectively. The repeat-mediated deletions cleanly removed the markers from transformed cpDNA without influencing the corresponding (homologous) endogenous genes (*rrn16* and *psbA*). However, it was not possible to control the recombination, and the desired deletion and/or marker-free plants were obtained by screening many seed progeny.

During development of chloroplast transformation in lettuce (*Lactuca sativa*), the cpDNA was transformed with *gfp* and *aadA* genes that had identical *rrn16* promoters (~200 bp); the genes were placed into the *trnA*/*trnI* intergenic region. Since the *gfp*

marker was flanked by the promoter repeat, it was spontaneously deleted in a small fraction of the progeny (Lelivelt *et al*. 2005).

In tobacco, Dufourmantel *et al*. (2007) introduced the herbicide-tolerance marker, HPPD (4-hydroxyphenylpyruvate dioxygenase), which confers tolerance to diketonitrile, into the plastid genome. However, the gene-fusion contained the 3'-UTR from *rbcL* (210 bp), and since it was introduced downstream of the endogenous *rbcL* gene, there was a background level of marker loss. Apparently this was due to recombination between the direct repeat of the *rbcL* 3'-UTR; the deletion was ~1.5 kb. They also used repeat-mediated deletion to remove the selecting *aadA* marker, which was used for cotransformation. In this experiment the HPPD marker was disrupted by an *aadA* expression cassette that was also flanked by a duplicated internal portion (403 bp) of HPPD. Recombination between the direct repeat of HPPD deleted the *aadA* cassette and restored HPPD-based herbicide tolerance.

Repeat-mediated deletion was also used to knock-out the *rbcL* gene (which encodes the large subunit of Rubisco) in tobacco (Kode *et al*. 2006). An *atpB*::*aadA* cassette formed a long direct repeat (649 bp) with the endogenous *atpB* gene. Since the *rbcL* gene was between the cassette and *atpB*, recombination between the *atpB* promoter repeats deleted *rbcL* and the *aadA* marker. Also of interest for the present study was the fact that Southern blotting of the *rbcL* knock-out line provided evidence of heteroplasmicity at the *rbcL* locus. However, sequencing of the remaining *rbcL* showed that it was a degenerate copy that most likely came from the nucleus. Nuclear DNA in

angiosperms, including tobacco, frequently contains cpDNA sequences (Ayliffe and Timmis 1992; Naciri and Manen 2010; Noutsos *et al*. 2005; Shahmuradov *et al*. 2003).

Despite the successes of repeat-mediated deletion, the method has limits: the length of the repeated sequence and distance between the repeat-pairs have to be gauged carefully, because recombination is strongly dependent on those parameters. Also, once the repeat is introduced into the plastid genome, there is no way to trigger, or prevent, the recombination from taking place. Thus, heteroplastomic plants will be obtained if the repeat-mediated deletion is lethal.

An alternative approach that was developed in tobacco by the Maliga lab is based on a site-specific phage recombinase (Thorpe *et al*. 2000; Thorpe and Smith 1998) that uses specific, introduced sequences to effect a recombinational deletion (Ow 2002). The recombinase from phage P1, called Cre, was used initially. Its substrate sequence, *loxP*, must flank the region to be deleted, and was introduced by chloroplast transformation (Corneille *et al*. 2001; Hajdukiewicz *et al*. 2001). Then, plastid-targeted Cre was introduced by crossing the transplastomic plant with a transgenic line that constitutively expressed modified Cre from the nucleus. This approach was used to demonstrate the essential nature of the plastid protease subunit, *clpP* (Kuroda and Maliga 2003). Subsequently, these investigators used transient transformation with plastid-targeted Cre to generate marker-free transplastomic plants (Lutz *et al*. 2006). However, the Cre-*lox* system was not without problems. The introduced substrate sequence is left behind after recombination; however, the principal obstacle was unwanted deletions in cpDNA promoted by Cre (Corneille *et al*. 2003; Hajdukiewicz *et al*. 2001).

Hence, a second-generation of this system was developed based on the phiC31 recombinase, Int (Kittiwongwattana *et al*. 2007). With this system, the recombinase recognition sequences (*attB*/*P*) that flank the plastid DNA to be deleted (a marker gene in this case) are not identical, thus decreasing the background recombination. The nucleus was transformed with a constitutively-expressed *Int* gene targeted to the chloroplast, and the marker gene was efficiently deleted. Fortunately, the Int recombinase does not recognize any naturally occurring sequences in tobacco cpDNA (Kittiwongwattana *et al*. 2007). However, like the Cre-*lox* system, a copy of the recombinase substrate is left behind.

The recombinase approach has the advantage of delaying the recombination event until the nuclear gene is introduced, but there is no control beyond that point. Also, there have not been other published reports of its use. A conditional or inducible system for cpDNA modification would have greater flexibility than the methods developed so far, but none has been reported. In this chapter, I describe such an approach in tobacco, which is also novel in that it is based on the homing endonuclease, I-*Cre*II. To make the modification of cpDNA inducible, a steroid-dependent expression system was used. This approach is seamless in that all of the introduced, non-native sequences are removed; plus, the deleted DNA is presumably destroyed in the process.

To accomplish this, I have taken advantage of the organelle's ability to repair a DSB in its genome using a SSA-like mechanism, as well as the general absence of long direct repeats in angiosperm cpDNA (Cosner *et al*. 1997; Jansen *et al*. 2006; Saski *et al*. 2005). DSB repair that proceeds by SSA between flanking direct repeats typically creates

a seamless deletion with concomitant destruction of the deleted DNA (Noskov *et al*. 2010; Odom *et al*. 2008). The DSB is created using plastid–targeted I-*Cre*II, which recognizes a sequence of 30-40 bp and specifically cuts the *psbA* gene in many plants, however, not in tobacco (Corina *et al*. 2009; Kim *et al*. 2005). Thus, the natural I-*Cre*II cleavage site was introduced using the chloroplast transformation plasmid, pMSK57 (Khan and Maliga 1999), which, by virtue of its 16S rRNA promoter, also creates a direct repeat of 84 bp with the endogenous *rrn16* gene (Lutz *et al*. 2007; Zoubenko *et al*. 1994). The endonuclease was introduced by crossing to a transgenic line that carried the chloroplast-targeted *I-CreII* gene (*rbcS:I-CreII*) under control of β-estradiol (Zuo *et al*. 2001). We propose to call this technique DREEM (for *d*irect *re*peat and *e*ndonuclease *m*ediated) modification of the plastid genome.

3.2 Results

*3.2.1 Creation of a transplastomic tobacco line with an I-*CreI*I cleavage site in the* MSK57 *background*

The chloroplast transformation plasmid, pMSK57 (Figure 11), contains a dual-function marker that is a fusion of *aadA*, which confers spectinomycin resistance, and *gfp*, which enables visual confirmation of transformed shoots by the GFP fluorescence (Khan and Maliga 1999). The *aadA:gfp* gene is located between *orf70B* and a valyl-tRNA gene (*trnV*) in the inverted repeat region of cpDNA. It was engineered for strong

---

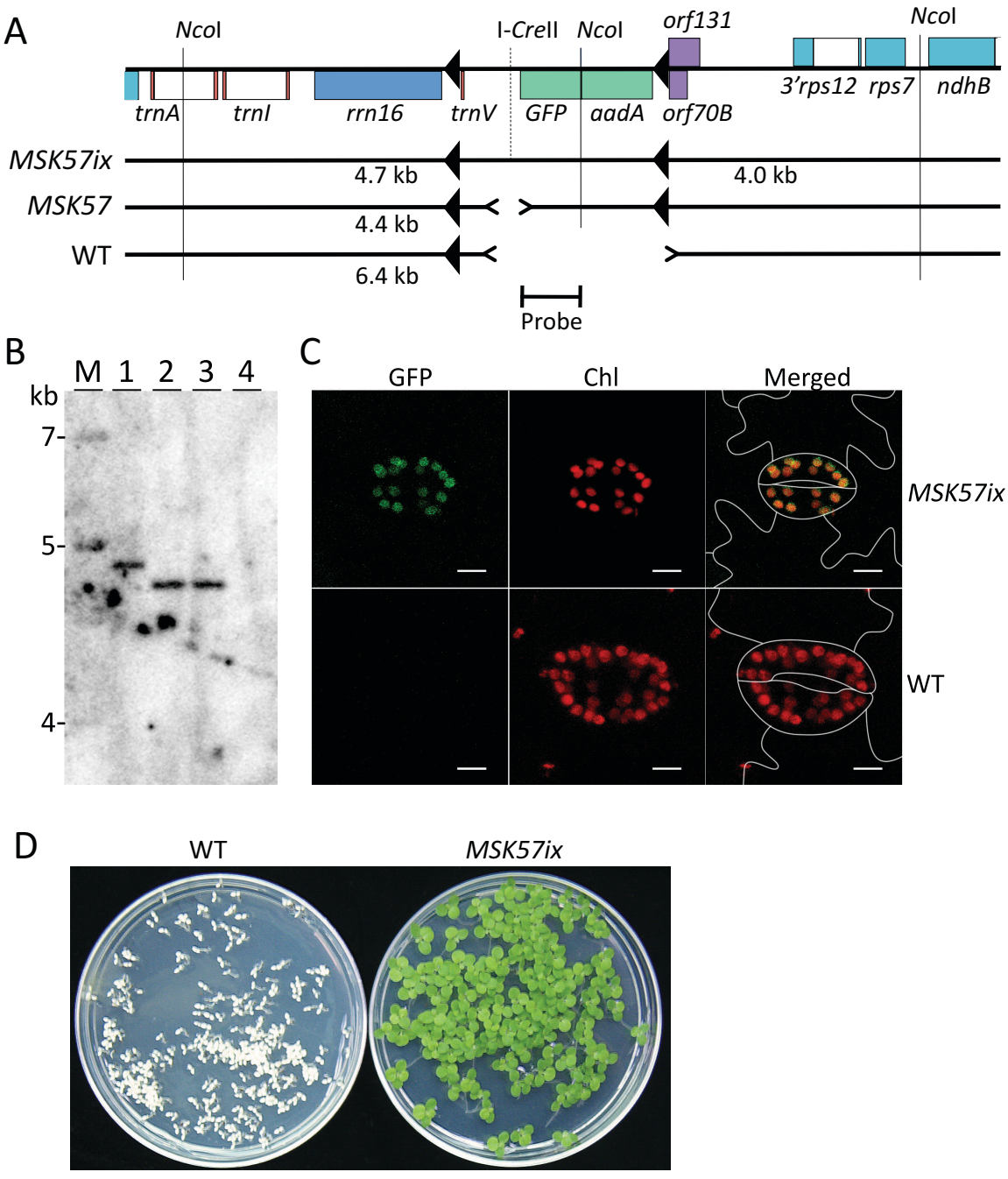
Figure 11. Insertion of an I-*Cre*I cleavage site, a selectable marker, and a flanking direct repeat into the tobacco chloroplast genome.

(**A**) Plasmid pMSK57ix was created by inserting an I-*Cre*I cleavage site in the 3' UTR of the *aadA:gfp* marker in the plastid transformation vector, pMSK57 (Khan and Maliga 1999). Shown is the plastid genomic map for *MSK57ix*, and below the differences with *MSK57* and wild type (WT) are indicated. Angle brackets indicate the regions that are lacking in the respective genomes. The *rrn16* promoter (*solid triangle*) on the *aadA:gfp* marker creates a direct repeat (84 bp) with the endogenous *rrn16* gene. Genes above the line are transcribed from left to right (→), and those below the line from right to left (←).

(**B**) Southern hybridization of the *MSK57ix* transplastomic line. Equal amount of genomic DNA was digested with *Nco*I, and the *gfp* gene was used as the probe (indicated in (**A**)). The lanes were: M, ^{32}P-labeled 1-kb DNA ladder; 1, *MSK57ix* transformant; 2 and 3, *MSK57* transformants; 4, wild type.

(**C**) Confocal microscopy of the *MSK57ix* chloroplast transformant. Images of the GFP and chlorophyll (red) fluorescence were obtained from guard cells. Cells were outlined in the merged images. Scale bars = 10 µm.

(**D**) Segregation analysis for the *aadA;gfp* marker, which confers spectinomycin resistance. Seeds from a cross of *MSK57ix* (T0, ♀) with wild type (♂) were germinated on MS medium containing spectinomycin (500 mg/L). Wild type tobacco seeds were analyzed as a control, which produced bleached seedlings on spectinomycin.


A
NcoI
I-CreII
NcoI
orf131
NcoI
trnA
trnI
rrn16
trnV
GFP
aadA
orf70B
3'rps12
rps7
ndhB
MSK57ix
4.7 kb
4.0 kb
MSK57
4.4 kb
WT
6.4 kb
Probe
B
M 1 2 3 4
kb
7-
5-
4-
C
GFP
Chl
Merged
MSK57ix
WT
D
WT
MSK57ix

transcription by adding 84 bp of the native 16S rRNA promoter (Khan and Maliga 1999; Zoubenko *et al*. 1994), which, because of its location and orientation, creates a direct repeat of 84 bp with the endogenous 16S rRNA promoter (leftward arrowheads in Figure 11A). I modified pMSK57 (kindly provided by P. Maliga, Rutgers Univ.) by inserting a small section of the *Chlamydomonas psbA* (intronless) gene that contains the native recognition sequence for I-*Cre*II, which is ~40 bp spanning the exon 4-exon 5 junction (Corina *et al*. 2009; Kim *et al*. 2005). The I-*Cre*II target sequence was inserted just downstream of the stop codon for *aadA:gfp* in the 3' UTR, which was derived from the tobacco *psbA* gene (Khan and Maliga 1999). It should be noted that the I-*Cre*II cleavage site in the new plasmid, pMSK57ix, is located between the direct repeat of the 16S rRNA promoter (Figure 11A).

Tobacco (*Nicotiana tabacum* cultivar Petit Havana) leaves were subjected to particle bombardment with plasmid DNA using an established protocol (Lutz *et al*. 2006), and then incubated for sterile culture on spectinomycin-agar plates as described (Lutz *et al*. 2006). After shoot regeneration on spectinomycin 3-4 times, I was able to obtain a single transplastomic line (with pMSK57ix) that carried the expected DNA as verified by Southern blotting (Figure 11B, Lane 1) and PCR (described below). Confocal fluorescence microscopy of the *MSK57ix* line confirmed that expression of the *aadA:gfp* marker was restricted to chloroplasts (Figure 11C, top panels).

Examining the transmission of the transformed cpDNA into the next generation provides a genetic assessment of the homoplasmicity/heteroplasmicity of the plastid genome. The transplastomic *MSK57ix* line was cross-pollinated with wild-type tobacco

and seed progeny were screened for resistance to spectinomycin. As shown in Figure 11D, all of the seed progeny from the *MSK57ix* line were resistant to a high level (500 mg/L) of spectinomycin, whereas the wild-type progeny (WT) were bleached and stunted. Screening of additional seed progeny of *MSK57ix* × WT (~550 seedlings) on spectinomycin plates, also yielded only healthy green plants like those in Figure 11D. These results indicated that *MSK57ix* was essentially homoplastomic.

3.2.2 Creation of transgenic tobacco with an inducible rbcS:I-CreII *gene in the nucleus*

The plan was to express I-*Cre*II as an inducible nuclear *transgene* that would be imported to the chloroplast and cleave the introduced target. The β-estradiol-inducible Cre-*loxP* system (pX6) that I used in *Arabidopsis* (Chapter 2) was good for preventing unscheduled expression of a toxic gene. However, in wild type tobacco, a basal level of

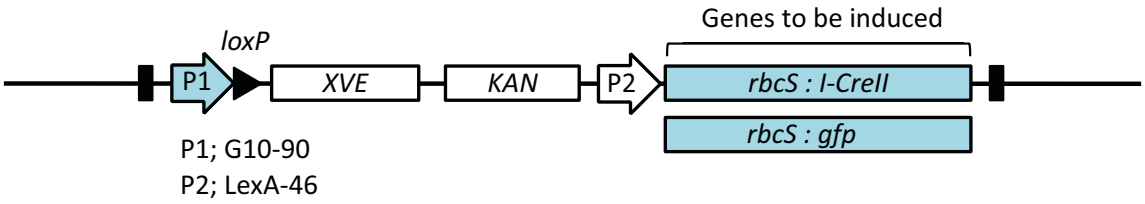


Figure 12. The β-estradiol controlled and chloroplast-targeted I-*Cre*II expression system.

The *Cre* gene and a *loxP* site were removed from plasmid pX6RIII (Figure 4), so that in these constructs, the XVE transcription factor, which is activated by β-estradiol, directly turns on the P2 promoter. With pRIII (top map), the fusion protein rbcS:I-*Cre*II is targeted to the chloroplast by the rbcS transit peptide. In the control plasmid pRGFP, the *gfp* gene replaced *I-CreII* (lower map). These constructs were introduced into the tobacco nuclear genome using *Agrobacterium*.

rbcS:I-CreII expression was expected to be innocuous, because the chloroplast genome lacks I-*Cre*II cleavage sites (Kim *et al*. 2005). Thus, the Cre-*loxP* recombination system (Figure 4) was removed from the pX6-based plasmid (Chapter 2), in order to make *rbcS:I-CreII* expression more responsive to β-estradiol, and ultimately reversible. The resulting construct, pRIII, is shown in Figure 12, along with the control construct, pRGFP, which had *gfp* instead of *I-CreII*. Both constructs were transformed into the wild-type tobacco cultivar (that was also used for chloroplast transformation) using *Agrobacterium* and selection on kanamycin (Zuo *et al*. 2001). The new transgenic lines, which were initially selected on the basis of a 3:1 segregation of kanamycin resistance in the T1 generation, were named for their transforming plasmids, *RIII* and *RGFP*, respectively. The growth of these lines, with or without β-estradiol, was indistinguishable from the parental wild type strain. Also, *rbcS:gfp* inducibility in the *RGFP* line was apparent to the eye under long-wave UV light, or with a fluorescence microscope (not shown).

3.2.3 Creation of the transplastomic-transgenic hybrids: stability of the spectinomycin resistance marker in the **rbcS:I-CreII** *background*

To create the final experimental lines, the transplastomic *MSK57ix* line was out-crossed to the transgenic *RIII* and *RGFP* lines. F2 progeny homozygous for the nuclear transgenes (either *rbcS:gfp* or *rbcS:I-CreII*) were germinated and examined for growth on spectinomycin (Figure 13A). All of the *MSK57ix*;*RGFP* progeny were completely resistant to the drug, similar to the *MSK57ix* parent. Most (>80%) of the *MSK57ix*;*RIII*

progeny were also resistant to spectinomycin; however, a fraction (2-8%) of these seedlings were highly susceptible (examples are shown in Figure 13B), and some others (2-11%) were partially resistant, as evidenced by their variegated (i.e., with both green and bleached sections) leaves (examples are shown in Figure 13B). The spectinomycin bleached sections of the variegated leaves presumably reflect clonal losses of spectinomycin-resistance during leaf development. It is noted that some of these plants produced fully green leaves at later times; this might be the result of a weakening of the selection due to falling spectinomycin concentrations in the medium, or the demand for cpDNA copies that confer spectinomycin resistance is less at later times (Rowan *et al*. 2004). Overall, these results indicate that the *aadA:gfp* locus is fairly, though not completely, stable in the nuclear *RIII* (*rbcS:I-CreII*) background. The minor fraction of spectinomycin-sensitive progeny obtained in the absence of the inducer probably reflects background expression of *rbcS:I-CreII* (which destabilizes the *aadA:gfp* locus as shown below).

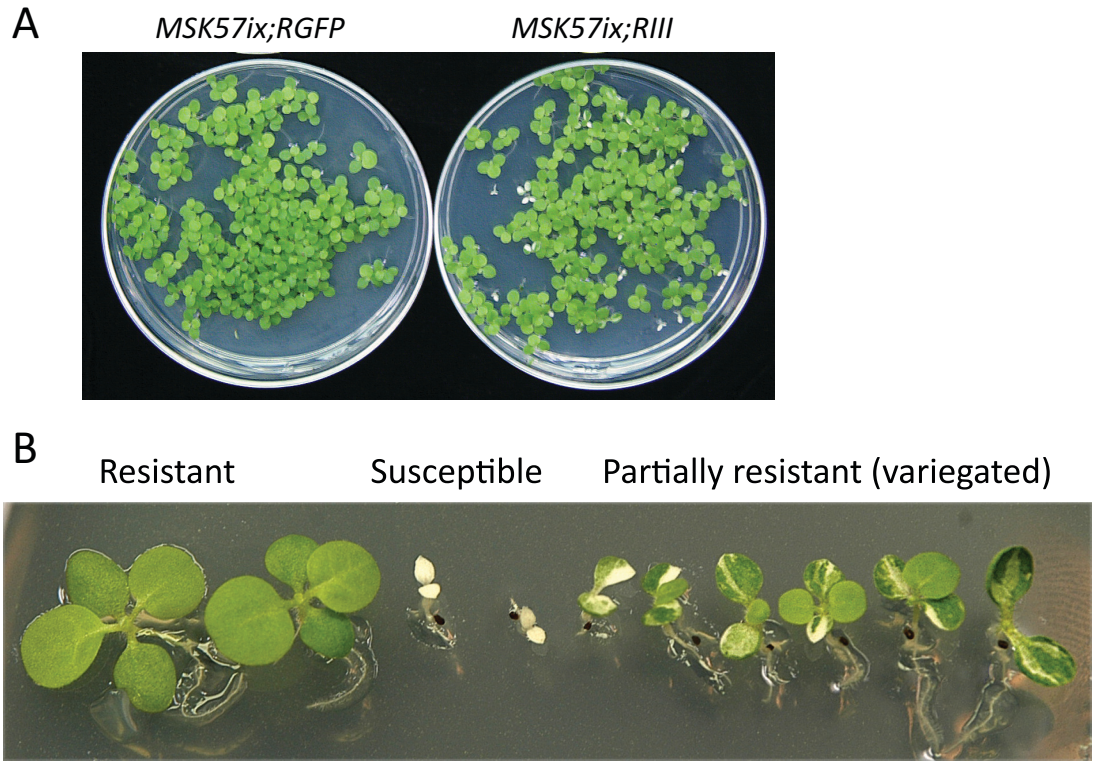


Figure 13. Stability of the *aadA:gfp* chloroplast marker in the *rbcS:I-CreII* nuclear background as assessed by spectinomycin resistance.

(**A**) The transplastomic line (♀, *MSK57ix*) was crossed with transgenic lines carrying either the *rbcS:I-CreII* gene (♂, *RIII*) or the *rbcS:gfp* gene (♂, *RGFP*). Seeds were germinated on MS medium containing spectinomycin (500 mg/L). Genotypes are indicated at the *Top*.

(**B**) Ten selected seedlings were transferred from the *MSK57ix;RIII* plate onto fresh MS + spectinomycin medium for further growth. The seedlings were selected as examples of spectinomycin resistant, susceptible and partially resistant (variegated) seedlings. Three of the variegated seedlings had produced green leaves by the time the photograph was taken.

3.2.4 Induction of rbcS:I-CreII *and repeat-mediated deletion of* aadA:gfp

To genetically examine the effect of β-estradiol on the *aadA:gfp* marker in the *MSK57ix:RIII* line. I tracked the loss of GFP fluorescence (Figure 14A), instead of the loss of spectinomycin resistance; this was done to avoid the selecting effect of the

antibiotic. With the *MSK57ix* chloroplast genome, the green GFP-based fluorescence is strong enough to mask much of the red fluorescence from chlorophyll, at least in young leaves. As shown in Figure 14B(a,b), the GFP fluorescence is stronger in the young emerging leaves than in the older leaves - the cotyledons and first true leaves barely fluoresced green while the young leaves closer to the shoot apex showed strong green fluorescence. When the GFP fluorescence was lost, as in Figure 14B(c,d), the entire shoot fluoresced red, regardless of age. Some variegated, with respect to GFP fluorescence, seedlings were seen in the absence of the inducer, and examples of these seedlings after 3 weeks growth are shown in Figure 14B (e,f).

Seeds of the *MSK57ix;RIII* and *MSK57ix* lines were germinated on MS-sucrose medium, with (10 μM) and without (0 μM) β-estradiol, and photographed at two weeks (Figure 14A). Essentially all of the seedlings from both lines were green and healthy, regardless of β-estradiol treatment. Also, the GFP fluorescence of the *MSK57ix* seedlings was unaffected by β-estradiol. However, all of the *MSK57ix;RIII* seedlings germinated on β-estradiol had lost visible GFP fluorescence from the cotyledons, which fluoresced red. In the absence of the inducer (0 μM β-estradiol), only ~10% of these seedlings showed evidence of a loss of GFP fluorescence (Figure 14A), which is consistent with the spectinomycin sensitivity assay (Figure 13). It should be noted that the *MSK57ix;RGFP* line could not be used as a control in this assay, because of its GFP fluorescence from the nuclear transgene, *rbcS;gfp*.

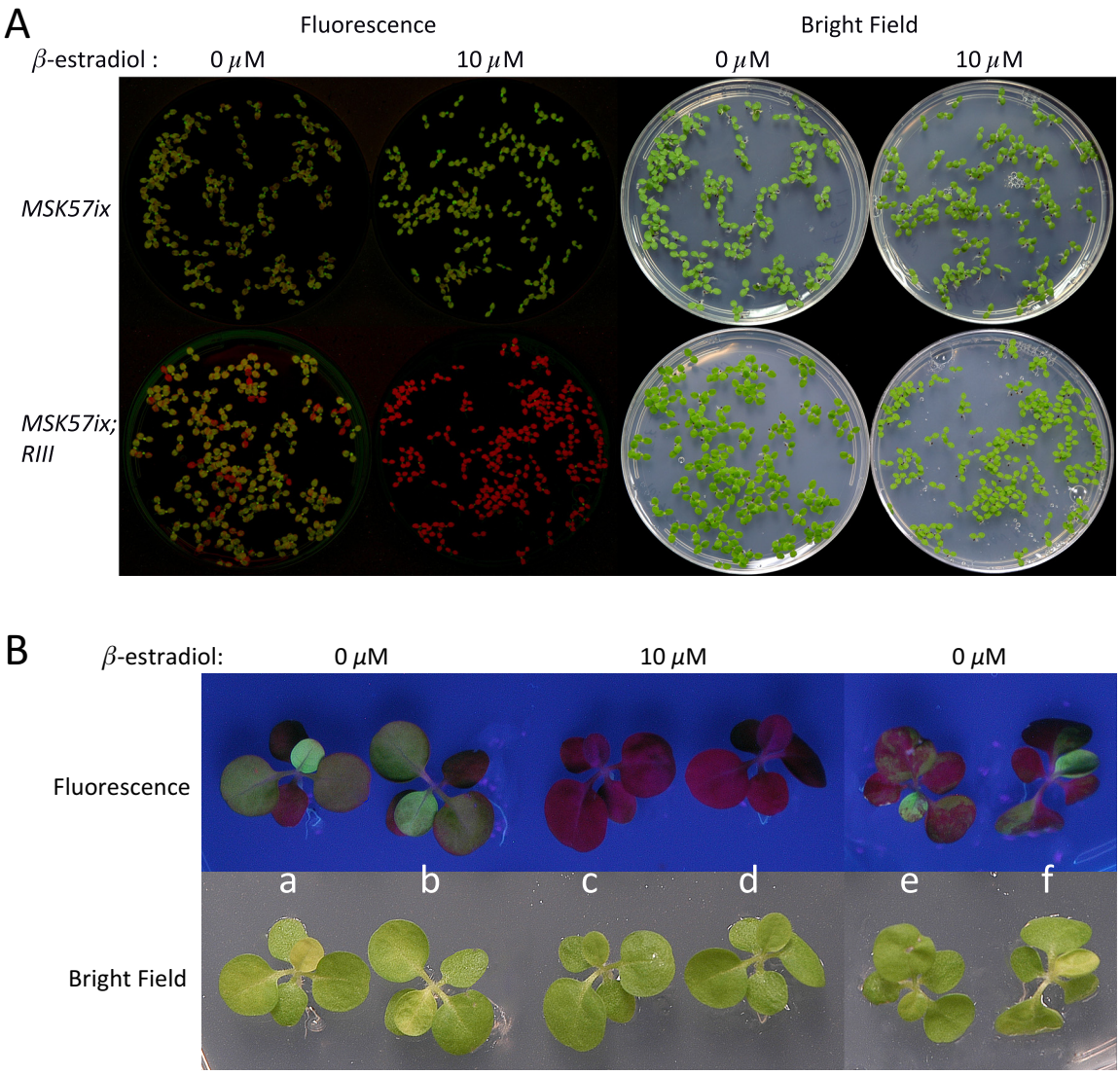


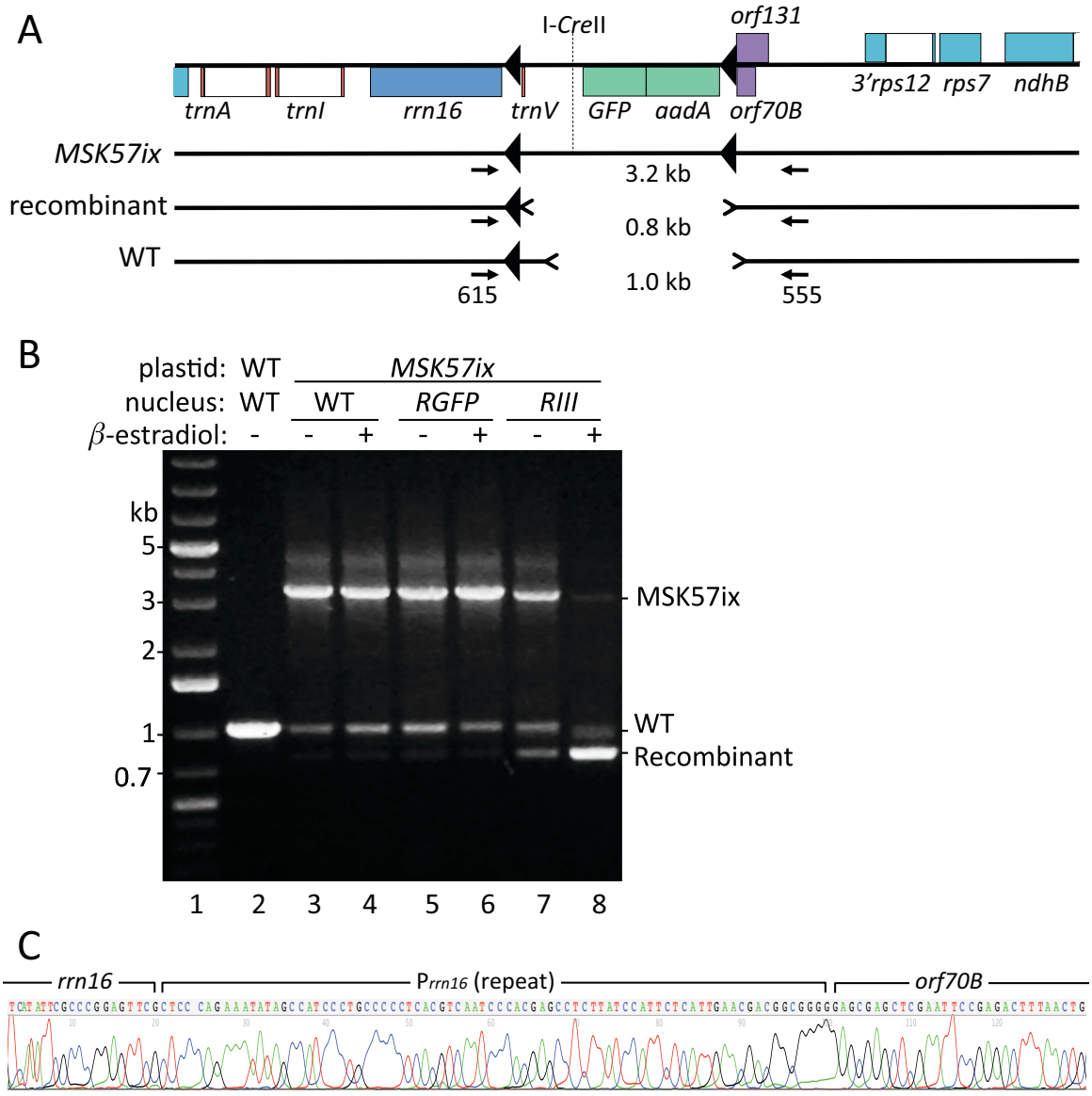
Figure 14. β-Estradiol induced loss of the *aadA:gfp* marker.

(**A**) Seeds from the *MSK57ix* and the *MSK57ix;RIII* (*rbcS:I-CreII*) lines were germinated on MS medium containing the indicated amount of β-estradiol. The plates were photographed 2 weeks after sown under long wave UV (for fluorescence) or white light. Plant genotypes are indicated to the left.

(**B**) Seedlings (3 weeks old) of the *MSK57ix;RIII* line were chosen to show the pattern of GFP expression, which depicted loss of the *aadA:gfp* gene. Seedlings were grown on media with the indicated concentration of β-estradiol.

PCR was used to physically examine the fate of cpDNA at the *aadA:gfp* locus (Figure 15). The map in Figure 15A is of the *MSK57ix* plastid genome. The change from wild-type cpDNA is indicated below (WT line), as is the difference with the recombinant cpDNA that is formed by SSA repair of the DSB using the repeats (leftward arrowheads). The PCR primers, which flank the inserted *aadA:gfp* transgene, and the expected sizes of the products are also indicated on the line diagrams.

Figure 15B shows an agarose gel of the PCR reactions performed with total DNA from wild-type plants (WT plastid and nucleus; lane 2), *MSK57ix* (lanes 3 and 4), *MSK57ix;RGFP* (lanes 5 and 6), and *MSK57ix;RIII* (lanes 7 and 8); also, the transplastomic and transplastomic-transgenic lines were grown with (+) or without (-) β-estradiol, as indicated. The gel shows that the 3.2-kb product from the *MSK57ix* plastome (lanes 3 and 4) is unaffected by β-estradiol treatment in the *rbcS:gfp* background (lane 6), but is substantially depleted in the *rbcS:I-CreII* background (lane 8). Moreover, there is a parallel increase in the 0.8-kb PCR product from the recombinant plastome in the latter plant (lane 8). The nature of the 0.8-kb PCR product was confirmed by direct DNA sequencing; the sequencer trace (Figure 15C) shows that the promoter of the *rrn16* gene had become precisely fused to the *orf70B* sequence. This is the predicted outcome of repairing the DSB induced by *rbcS:I-CreII* by SSA between the 16S rRNA promoter sequences.


A
I-CreII
orf131
trnA
trnI
rrn16
trnV
GFP
aadA
orf70B
3'rps12
rps7
ndhB
MSK57ix
recombinant
WT
3.2 kb
0.8 kb
1.0 kb
615
555
B
plastid: WT MSK57ix
nucleus: WT WT RGFP RIII
β-estradiol: - - + - + - +
kb
5
3
2
1
0.7
MSK57ix
WT
Recombinant
1 2 3 4 5 6 7 8
C
rrn16
Prrn16 (repeat)
orf70B

The gel also shows that there is a significant level of the recombinant plastome (0.8-kb PCR product) in the DNA from the *MSK57ix;RIII* plants that were not treated with β-estradiol (lane 7), which is consistent with the phenotypic assays (Figures 13 and 14). Thus, this data also confirms that the minor fraction of plants with the *rbcS:I-CreII* nuclear background that lost spectinomycin resistance and GFP fluorescence in the absence of inducer were losing the *aadA:gfp* marker because of recombinational DSB repair.

Although difficult to see in photographs, the original gels also showed evidence of a very low level of the recombinant plastome in plants lacking the *rbcS:I-CreII* transgene (Figure 15B, lanes 3-6). Assuming that this faint 0.8-kb PCR product is what it seems, the data suggests a low level of spontaneous recombination occurs between the 84-bp repeats (Lutz *et al*. 2007), which are longer than those typically found in tobacco cpDNA. It should be noted, however, that this level of recombinant cpDNA was not high enough to produce evidence of spectinomycin sensitivity in the seedling assay (Figure 13). This

---

Figure 15. PCR analysis at the *aadA:gfp* locus of cpDNA.

(**A**) Map of *MSK57ix* cpDNA at the *aadA:gfp* locus, and resulting PCR products. The *rrn16* promoter that forms the direct repeat is indicated by the black arrowheads. The portions of cpDNA lacking from the recombinant and wild-type cpDNAs are indicated below the map, as are the sizes of the PCR products obtained for each with the indicated primers. The other labels are the same as in Figure 11.

(**B**) Agarose gel of the PCR products. Equal DNA from tobacco seedlings germinated with (+) or without (-) β-estradiol (as indicated above the lanes) was subjected to PCR with primers 555 and 615. Also indicated are the nuclear (WT, wild type; RGFP, *rbcS:gfp*; RIII, *rbcS:I-CreII*) and plastid (WT and *MSK57ix*) genotypes. Size markers (lane 1) were the 1-kb Plus GeneRuler (Fermentas).

(**C**) Sequence trace of the 0.8-kb PCR product from *MSK57ix;RIII* treated with β-estradiol. The *rrn16* gene and its promoter are fused to *orf70B*.

is an important point because of the wild-type (1 kb) PCR product that was obtained in all the transplastomic (*MSK57ix*) plant lines (lanes 3-8). This product is most likely amplified from cpDNA sequences in the nucleus (Ayliffe and Timmis 1992). If it originated from wild-type copies of plastid DNA, there should have been some evidence of spectinomycin sensitivity, or loss of GFP fluorescence, in the seedling assays (Figures 11-14) like there was for the *MSK57ix;RIII* line (Figures 13 and 14), which had a comparable level of PCR product (0.8 kb) derived from recombined plastid genome.

3.2.5 Deletion of a 3.8-kb segment of cpDNA including ~60% of ycf1

In order to further demonstrate the utility of the DREEM technique, I generated tobacco strains for the inducible deletion of a sizable segment of native cpDNA. *ycf1* is one of the two largest conserved ORFs (1901 amino acids in tobacco) in cpDNA, and its function is unknown. A chloroplast transformation plasmid (pOT43) was created with the structure shown in Figure 16, and used to transform wild-type tobacco as described above

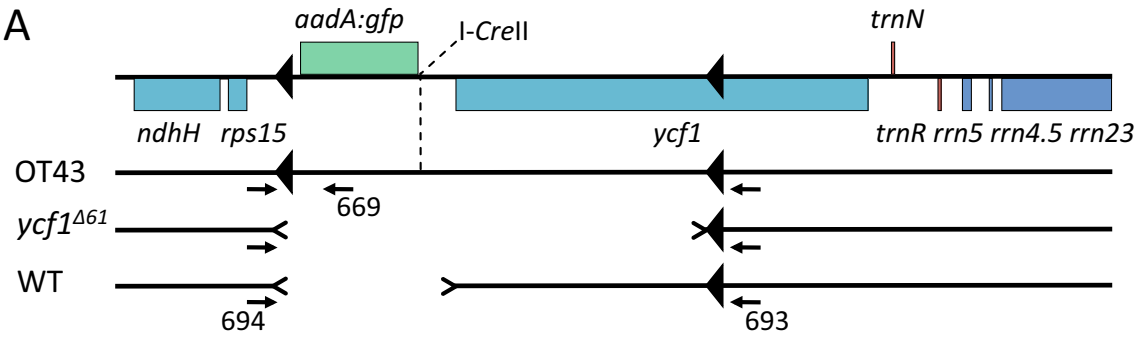
---

Figure 16. PCR analysis of plastid DNA in 3 progeny from the ♀ *OT43* × ♂ *RIII* cross that were germinated on β-estradiol.

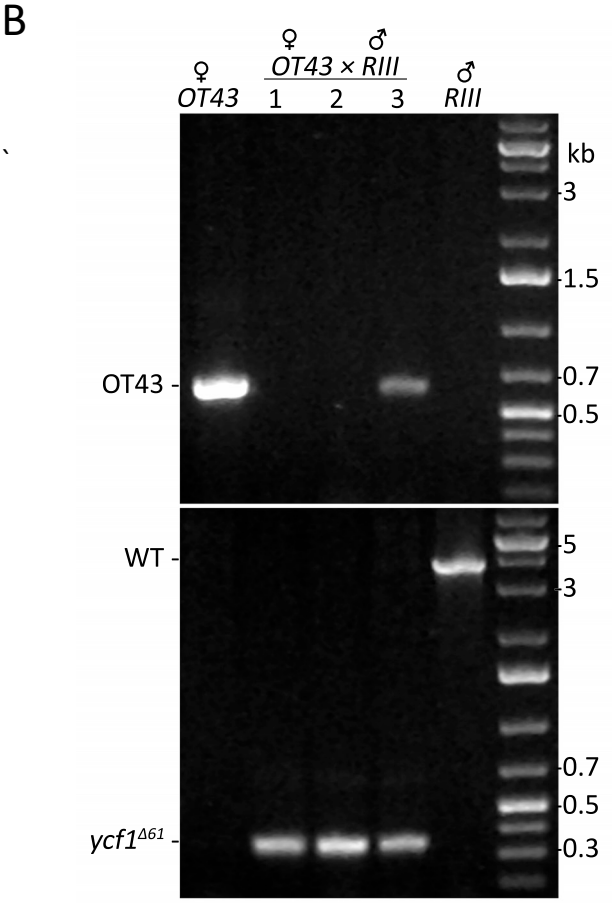
(**A**) Map of cpDNA at the *ycf1* locus in the transplastomic line, *OT43.* Genes above the line are transcribed from left to right, whereas those below it are transcribed from right to left. The large arrowheads mark the 227-bp direct repeat of *ycf1*. The lines below the map indicate regions that are not found in wild-type plastid DNA (WT), or are deleted by SSA-like repair of *I-CreII* cleavage; since the deletion removes 61% of the *ycf1* ORF, the recombined plastid DNA was named $ycf1^{\Delta 61}$.

(**B**) F1 progeny of ♀ *OT43* × ♂ *RIII* were germinated on an induction medium. Leaves from three different seedlings were analyzed by PCR using OT43-specific primers 694/669 (*upper panel*) and direct repeat flanking primers 694/693 (*lower panel*). Progeny of parental *OT43* and *RIII* lines were included as control. The expected size (kb) of each PCR product is indicated.

A



B



Plant	Primers 694/669
OT43	0.6
ycf1$^{\Delta 61}$	-
WT	-

Plant	Primers 694/693
OT43	6.3
ycf1$^{\Delta 61}$	0.3
WT	3.8

for *MSK57ix*. The resulting transplastomic line, *OT43,* has a direct repeat (227 bp) of an internal portion of *ycf1* situated downstream of the endogenous gene, an *aadA:gfp* marker, and an I-*Cre*II cleavage site (Figure 16A). Recombinational repair of the DSB by SSA between the direct repeats would delete ~6 kb of DNA total, and ~3.5 kb, or 61%, of *ycf1*. It also creates a stop codon for the truncated ORF, $ycf1^{\Delta 61}$, that is in the intergenic region, 99 bp from the *ycf1* sequence; thus adding 33 amino acids to the C-terminus. The OT43 line was healthy and otherwise resembled wild-type tobacco (not shown).

To analyze the effect of the endonuclease (*rbcS:I-CreII*), *OT43* was crossed with *RIII* and seed progeny were germinated on induction medium. Several plants showed evidence of reduced growth and pale virescent leaves, and 3 of them were subjected to PCR analysis (Figure 16B). Two of the progeny plants (*1* and *2*) gave no PCR product with the *OT43*-specific primer pair (*top gel*), however, the third progeny plant (*3*) did give a band of 0.6 kb, which is the expected size for unrecombined *OT43* DNA. The parental *OT43* and *RIII* lines (*outer lanes*) were analyzed in parallel, and gave positive and negative signals, respectively, with the 694/669 primer pair. PCR was also performed with primers flanking the direct repeat (*lower gel* in Figure 16B), and a prominent band of ~0.3 kb, which is close to the size expected for the deleted genome ($ycf1^{\Delta 61}$), was obtained with all 3 progeny plants. This product was also sequenced which confirmed its identity. For the parental lines, the gel shows the expected 3.8-kb product from *RIII* (which has a wild-type plastid), but the 6.3-kb product expected for *OT43* was apparently too long to amplify under these conditions; thus, the reason for the *OT43*-specific primers (*top gel*). These results show that a deletion of 3.5 kb of native cpDNA (plus the

introduced sequence) was obtained as predicted, and with a reasonably high efficiency in 2 of the 3 lines examined.

3.3 Discussion

I have developed a new approach for modifying the plastid genome, which we have named DREEM (for *d*irect *re*peat and *e*ndonuclease *m*ediated). An analogous method for modifying the yeast nuclear genome, called TREC (for *t*andem *r*epeat coupled with *e*ndonuclease *c*leavage), was recently reported (Noskov *et al*. 2010). Unlike TREC, which employs a widely-used yeast endonuclease (I-*Sce*I), in DREEM we use I-*Cre*II, which is a newer enzyme derived from a *Chlamydomonas* chloroplast intron, *Cr.psbA4*. I-*Cre*II cleaves intron-minus *psbA* from a number of plants, but not tobacco (Kim *et al*. 2005). This was fortuitous, however, because chloroplast transformation in tobacco is reliable; thus, we could introduce the I-*Cre*II cleavage site as it was needed. In addition to the cleavage site, the transforming DNA contains a modestly-sized sequence (80-200 bp) that, after integration, forms a direct repeat with an endogenous sequence; this repeat defines the subsequent deletion. Since the direct repeat flanks the endonuclease cleavage site, repair of the DSB by SSA (between the repeats) deletes the intervening DNA and one of the repeat copies. The deletion also removes the I-*Cre*II target site, thereby stabilizing the DNA against further cleavage.

DREEM modification has several advantages: (*i*) it is seamless, in that foreign, introduced DNA is not left behind as it is with the recombinase-mediated methods (e.g., with the Cre-*lox* system, a *loxP* site remains in the genome after recombination); (*ii*) it is

under chemical control, by virtue of the β-estradiol receptor-activator for *rbcS:I-CreII* expression – this feature should facilitate phenotypic studies of lethal mutations; (*iii*) it is immediate in that the deleted DNA is presumably destroyed during the SSA process – with recombinases the deleted DNA is probably lost gradually by sorting out during subsequent plastid and cell divisions (Staub and Maliga 1994); (*iv*) the nuclear *transgenes* can be subsequently removed in a single cross, and (*v*) it can be used to make other types of gene modifications, such as point mutations or fusions.

I took advantage of a chloroplast transformation construct developed previously for tobacco, pMSK57, which uses the *rrn16* promoter to drive expression of the *aadA:gfp* selection marker (Khan and Maliga 1999). After inserting the native target for I-*Cre*II at the 3' end of *aadA:gfp* and transformation into the plastid, a direct repeat was created with the endogenous 16S rRNA promoter that flanks the I-*Cre*II cleavage site. The integrated DNA was inherently quite stable, partly because of the modest size of the repeated sequence, and the fact that the repeats were 2.5 kb apart. In the *rbcS:I-CreII* nuclear background, recombination between the promoter repeats was detectable even without the inducer, affecting 3-10% of the seedlings in the spectinomycin-resistance assay. This was most likely due to a low basal expression of *rbcS:I-CreII* in the absence of β-estradiol. With the inducer in the medium, however, recombinational deletion of the *aadA:gfp* marker approached 100%, and no other effects on the phenotype of the plants were observed. Thus, based on the marker experiment, DREEM modification of the plastid genome is highly efficient.

In some situations, a lower basal level of recombination may be needed. This could be achieved by using the intact pX6 system, instead of this modified version, to control *rbcS:I-CreII* expression. With pX6, target gene expression is β-estradiol activated, but also dependent on Cre-*lox* recombination (Zuo *et al*. 2001). I used that system with *Arabidopsis* (Chapter 2), because its *psbA* gene is cleaved by I-*Cre*II. Alternatively, *rbcS:I-CreII* could be introduced into the transplastomic line by nuclear transformation, instead of crossing, and transgenic lines could be screened for very low basal excision rates (of the *aadA:gfp* marker) by plating on spectinomycin without the inducer.

In *Arabidopsis*, the DSB created by I-*Cre*II was repaired with a number of junction sequences, most of which involved microhomology (6 -12 bp perfect repeats and 13-16 bp imperfect repeats). In tobacco, however, it was repaired mainly with the introduced repeat, which was 84 bp for *MSK57ix* and 227 bp for *OT43*. It should be said that, although the induced deletion in *OT43* (*ycf1*$^{\Delta 61}$) was much larger, 6 kb versus 2.5 kb, it is not clear that the longer repeat was really necessary; an 84-bp repeat may have worked just as well. If we examine plastid DNA around the I-*Cre*II site in *MSK57ix* for direct repeats that might compete with the *rrn16* promoter repeat, there are none that come close to 84 bp, and all the perfect repeats are 12 bp or smaller (Figure 17). This suggests that if a much longer repeat is available in the vicinity of a DSB, it is preferentially used for repair, at least by these pathways.

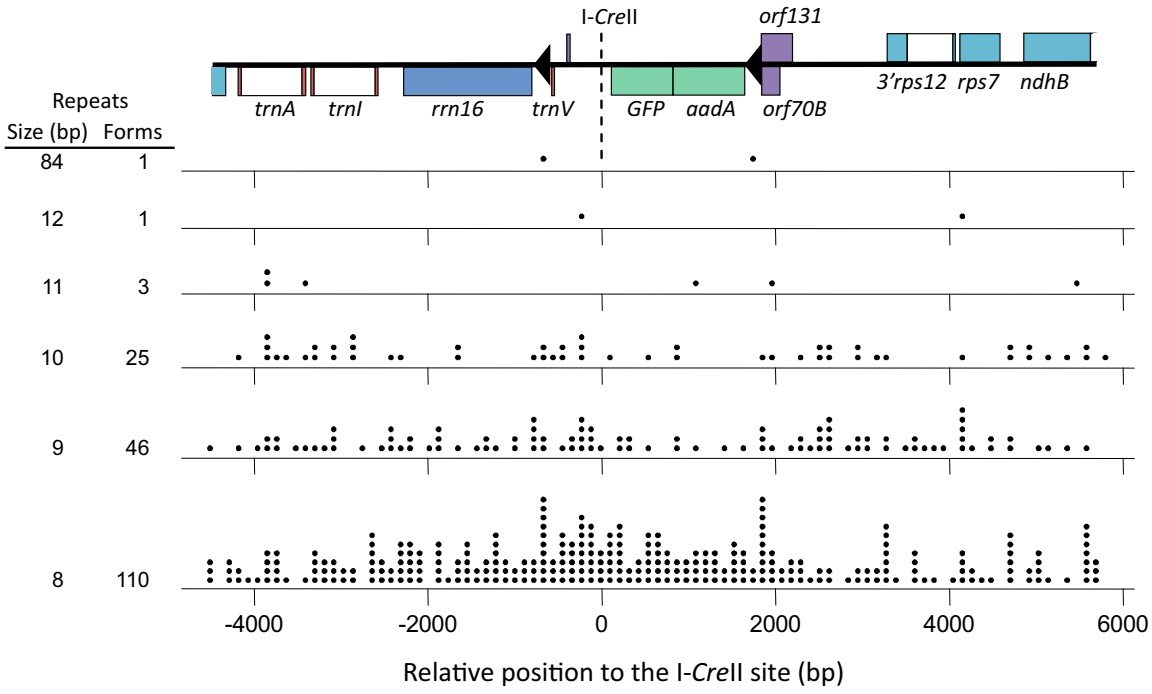


Figure 17. Distribution of direct repeats flanking the I-*Cre*II cleavage site in the *MSK57ix* cpDNA. Each dot represents a direct repeat element of the indicated size. Relative distance (bp) to the I-*Cre*II site is shown on the horizontal axis. The numbers of different repeats of same size (Forms) are indicated.

It may also be noteworthy that the *trnV*(GAC) gene, one of the two valyl-tRNA genes in the tobacco chloroplast, was deleted from the *MSK57ix* plastid along with the *aadA:gfp* marker (Figure 15). This seemed not to have any effect on the plant's growth and development, indicating that it is not an essential tRNA gene. The chloroplast genomes of the maiden hair fern, *Adiantum capillus-veneris* (Wolf *et al*. 2003), and the protist euglenoid, *Euglena gracilis* (Hallick *et al*. 1993), lack this valyl-tRNA, and have only the *trnV*(UAC) gene. If a "2 out of 3" pairing rule is invoked, the *trnV*(UAC) valyl-tRNA could read all four valine codons: GUU, GUA, GUC, and GUG (Pfitzinger *et al*.

1990). Valine codon usage by chloroplasts is skewed; in tobacco, the combined frequency of the GUU and GUA codons is 77%, whereas the GUC and GUG codons are used less frequently (23%) (Nakamura *et al*. 2000; Shimda and Sugiuro 1991). The importance of biased codon usage in tobacco chloroplasts is not yet clear, but it seems to be less critical than in *Chlamydomonas* (Franklin et al. 2002).

It has been of great commercial and public interest to have control of genetically modified organisms (GMO) used in agriculture, energy and medicine (Haslberger 2000; Roff 2009). The chloroplast genome has been an attractive target for genetic engineers, in part, because it offers better containment of foreign genes (Friesen *et al*. 2003; Warwick *et al*. 2009). The recent demonstration that there is leakage in the uniparental (maternal) inheritance of cpDNA, such that chloroplast *transgenes* can be transmitted through pollen has raised new concerns (Ruf *et al*. 2007; Svab and Maliga 2007). Removing as much foreign DNA as possible from the GMO's plastome could help with this issue (Kittiwongwattana *et al*. 2007). The features of DREEM modification mentioned above are highly compatible with this goal.

Chapter 4. Materials and methods

4.1 Plant and bacterial strains

Arabidopis thaliana Columbia (Col-0), *Nicotiana tabacum* cultivar Petit Havana and transgenic plants were grown either axenically on Murashige-Skoog (MS) medium (Murashige and Skoog 1962) or in soil. *Arabidposis* was grown at 23 °C under continuous light (150 μmol photons $\cdot$ m^{-2} $\cdot$ sec^{-1}) and tobacco was grown at 28 °C under a 16 hrs light / 8 hrs dark photoperiod, unless stated otherwise. *Agrobacterium tumefaciens* GV3101 (pMP90) was used for plant transformation. *Escherichia coli* DH5α was used for plasmid construction and propagation. *E. coli* Rosetta pLysS-RARE (Novagen) was used to express recombinant proteins and assay I-*Cre*II activity *in vitro* and *in vivo*.

4.2 Plasmids for chloroplast-targeted I-*Cre*II expression

The N-terminal 59 amino acids of *rbcS1A* (AT1G67090) was amplified using oligos 485 and 486 (Table 2), which have *Xho*I and *Afe*I sites, from TAIR U13397 (Krebbers *et al*. 1988) obtained from the Arabidopsis Biological Resource Center (The Ohio State University). *I-CreII* was amplified from pI-CreII (Kim *et al*. 2005) using 488 and 489 (Table 2), which removed the stop codon and added *Afe*I and *Avr*II sites. GFP was amplified from pX6GFP (Zuo *et al*. 2001) using oligos 490 and 491 (Table 2), which added *Avr*II and *Spe*I sites. The three products were cloned into the *Sma*I site of pUC18, creating pUC18-rbc, -ICre, and -GFP. GFP was excised with *Avr*II and cloned into *Avr*II-linearized pUC18-ICre, creating pUC18IIIG. The rbcS fragment was excised from

pUC18-rbc with *Bam*HI and *Afe*I, and cloned into *Bam*HI-*Afe*I digested pUC18IIIG, creating pUC18RIIIG. Its insert (*rbcS:I-CreII:gfp*) was completely sequenced, excised with *Xho*I and *Spe*I, and cloned into *Xho*I-*Spe*I cut pX6GFP; the final plasmid was pX6RIIIG. GFP was removed from pX6RIIIG by re-circularizing after cutting with *AvrI*I and *Spe*I, producing pX6RIII. The Cre-*lox*P system was removed from pX6RIII by re-circularizing after digstion with *Xho*I and *Sal*I (partially); the resulting plasmid was pRIII.

4.3 Plasmids to verify I-*Cre*II activity of fusion proteins

To test activity of the I-*Cre*II fusion protein, the pRSETA expression vector (Invitrogen), which confers ampicillin resistance on the host cell, was used. The insert of pUC18RIIIG (*rbcS:I-CreII:gfp*) was excised with *Xho*I and *Eco*RI, and cloned into *Xho*I-*Eco*RI-digested pRSETA, creating pRSET-RIIIG. Then, pRSET-RIIIG was cut with *Avr*II and re-circularized, thereby removing the GFP portion of the transgene. This plasmid was called pRSET-RIII.

Two I-*Cre*II substrate plasmids and a corresponding control plasmid were created as follows. Plasmid pE4E5, which contains the I-*Cre*II cleavage site as a fusion of exons 4 and 5 of *psbA* (Kim *et al*. 2005), was cut with *Dra*I and re-circularized, in order to remove the ampicillin resistance gene. The new plasmid which retains the kanamycin resistance marker, was named pDrail. To create a plasmid with multiple I-*Cre*II cleavage sites, the insert from pDrail was excised with *Spe*I and *Xba*I, and then ligated back to pDrail that had been cut with the same enzymes and dephosphorylated. A plasmid with five I-*Cre*II sites in tandem was found, and named pDrais. A control plasmid with no I-

*Cre*II cleavage site was created from pDrail by digesting it with *Eco*RI, followed by self-ligation; this plasmid was called pRey.

4.4 Plasmids for plastid gene deletion

A plastid transformation plasmid with an I-*Cre*II cleavage site was created as follows. The *Hin*dIII-*Sac*I fragment (49 bp) of pUC18 was ligated with pE4E5 (Kim *et al*. 2005) that had been cut with the same enzymes, creating pE4E5X. The insert of pE4E5X was excised with *Xba*I, and cloned into the *Xba*I site of the plastid transformation vector pMSK57 (Khan and Maliga 1999); the new plasmid was called pMSK57ix.

For deletion mutagenesis of the plastid *ycf1* gene, a plasmid, pOT43, was designed to introduce an I-*Cre*II cleavage site flanked by a direct repeat of part of *ycf1*. To this end, a series of intermediate plasmids were created as follow. Tobacco chloroplast DNA (1 μg) was digested with *Eco*RV and separated on an agarose gel. The 5-kb fragment (GenBank NC_001879: nt 123964-129280) was purified and cloned into the *Sma*I-linearized pUC19 vector, which produced pOT24. The insert of pOT24 was excised with *Kpn*I and *Bam*HI and then ligated with pDrail that was cut by the same enzymes, resulting in pOT32. Phosphorylated oligonucleotides 658 and 659 (Table 2) were annealed and ligated with *Xba*I-linearized pMSK57, which created an I-*Cre*II cleavage site and gave rise to pOT35. To create direct repeats, an internal *ycf1*-DNA segment (0.2 kb) was amplified using tobacco chloroplast DNA as template and primers 672 and 673 (Table 2). The PCR product was cleaved with *Eco*RI, and ligated to *Eco*RV-*Eco*RI-cut

pOT35, creating pOT37. The insert of pOT37 was excised with *Eco*RV and *Stu*I, and ligated to *Sca*I-cut pOT32; the final plasmid was pOT43.

4.5 Plasmid exclusion assay for I-*Cre*II activity

The *E. coli* Rosetta pLysS-RARE strain (Novagen) was used for the plasmid exclusion assay. It was transformed with one of the I-*Cre*II substrate plasmids (pDrail, pDrais, or pRey) and selected on kanamycin. Then, these lines were transformed with the I-*Cre*II fusion-protein plasmids (pRSET-RIIIG, pRSET-RIII, and the vector), and aliquots were plated on LB medium containing either: kanamycin (34 μg/ml), ampicillin (50 μg/ml), or both antibiotics. Since the substrate plasmids confer kanamycin resistance, and the enzyme plasmids confer ampicillin resistance, transformants that are resistant to both antibiotics have both plasmids (i.e., they are double-transformants). Colonies were counted after 18 hrs at 37°C, and the survival rate of double transformants calculated as follow; Survival rate (%) = (cfu μg^{-1} DNA on ampicillin and kanamycin) / (cfu μg^{-1} DNA on ampicillin) × 100.

4.6 *In vitro* I-*Cre*II assay using bacterial extracts

E. coli Rosetta pLysS-RARE strains (50 ml each) carrying plasmid pRSET-RIIIG, pRSET-RIII, or pRSETA (Vector) were grown until the A_{600} reached 0.5. Then, isopropyl-β-D-thiogalactopyranoside was added to 1 mM, and the cells were incubated for 2 hrs at 30 °C. After harvesting by centrifugation, the cells were resuspended in 5 ml of ice-cold 10 mM Tris-HCl, 0.1 mM EDTA pH 8.0, lysed by sonication, and centrifuged

for 10 min at 13,000 × g (4°C). The supernatant of this spin contained endonuclease activity, which was assayed by adding 5 μL of extract to a 25-μL reaction mixture that also contained 20 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, and 500 ng of substrate plasmid. The mixtures were incubated at 37°C for 30 min, and then analyzed by electrophoresis in 1%-agarose gels containing ethidium bromide.

4.7 Plastid transformation

Tobacco chloroplasts were transformed by particle bombardment (Lutz *et al*. 2006). Leaves from greenhouse-grown plants (3-8 weeks old) were surface-sterilized with 70% ethanol and 0.05% sodium hypochlorite, and bombarded with DNA-coated gold particles (1 μg DNA/5 mg gold, ø = 1 μm, kept on ice until used) using the helium-driven PDS-1000 (BioRad). Rupture discs of 1,100 psi were used under vacuum of 28 in. Hg, and the leaf tissue was 9 cm from the stopping plate. Leaf discs (1 cm × 1 cm) were placed abaxial-side up and incubated on RMOP media (Cséplő and Maliga 1984) containing spectinomycin (500 mg/L) up to 12 weeks at 28 °C under a 16 hrs light/ 8 hrs dark photoperiod. Green shoots were visually inspected for GFP fluorescence under long wave UV. Green fluorescent tissues were excised and subjected to regeneration 2~3 times to obtain shoots that were completely green fluorescent. The shoots were transferred to and cultivated on RM medium (Murashige and Skoog 1962) for 2-4 weeks, and the regenerated plantlets were potted and grown in the greenhouse.

4.8 Plant nuclear transformation

Arabidopsis was transformed by the floral dip method (Clough and Bent 1998), and seeds were germinated on MS medium contaning kanamycin (50 mg/L). Tobacco was transformed by the leaf disc method (Rogers *et al*. 1986). Briefly, fully expanded leaves were taken from tobacco plants that were cultivated in a greenhouse, and surface-sterilized with 70% ethanol and 0.05% sodium hypochlorite. Leaf discs (1 cm × 1 cm square) were prepared and co-cultured with *Agrobacterium* (carrying the transformation plasmid) on RMOP agar medium (Cséplő and Maliga 1984) for 2 days, then on RMOP medium containing carbenicillin (500 mg/L) and kanamycin (50 mg/L) for 7 days, and finally on RMOP medium with kanamycin alone for 4-8 weeks. Green shoots were transferred to and cultured on RM medium (Murashige and Skoog 1962) for 2-4 weeks, and the regenerated plants were transplanted into soil and grown in the greenhouse.

4.9 Cross-breeding to introduce rbcS:I-*Cre*II into the transplastomic tobacco

The transplastomic lines (*MSK57*, *MSK57ix*, and *OT43*) contain the *aadA:gfp* marker which confers spectinomycin resistance and GFP fluorescence. Seeds (100-200) were examined on MS medium containing spectinomycin (500 mg/L) every generation (T1-T3 generation). All of the progeny (> 500 total) were spectinomycin resistant and GFP positive; such transplastomic lines were considered as homoplastomic and used for the subsequent cross.

The transgenic (nuclear) tobacco lines (*RIII* and *RGFP*) contain a kanamycin resistant marker. These lines (T1 generation) that showed Mendelian segregation (3:1) for kanamycin resistance were considered to have a single T-DNA insertion. Transgenic lines homozygous for the kanamycin resistance marker were indentified by analyzing progeny on a MS-kanamycin (50 mg/L) medium in the next generation (T2) of self-pollinated plants.

To perform the crosses, the transplastomic lines (*MSK57*, *MSK57ix*, and *OT43*) were pollinated with the respective homozygous transgenic lines (*RIII* and *RGFP*). The F1 hybrids were heterozygous for the kanamycin marker. In the F2 generation of self-pollinated plants, homozygotes for the kanamycin marker were identified by analyzing progeny on MS-kanamycin medium. Both F1 (heterozygotes) and F2 (homozygotes) hybrids were germinated on MS medium containing 10 μM β-estradiol to induce expression of *rbcS:I-CreII* or *rbcS:gfp*; there was no difference between these lines in the effect of the steroid, which was expected from the dominant nature of the transgenes (*rbcS:I-CreII* and *rbcS:gfp*).

4.10 RT-PCR

Total RNA was isolated from 2-week old *Arabidopsis* (~50 mg fresh weight) using Trizol (Invitrogen) (Chomczynski and Mackey 1995) and reverse transcribed at 42°C for 50 min (1 μg RNA in 25 μL) with Superscript II and oligo-dT_{18}. Aliquots (0.05 vol) were subjected to PCR with primers specific for *I-CreII* (488 and 489) and *Act2* (535 and 536), and separated on 1% agarose gels for digital imaging.

4.11 Analyses of *Arabidopsis* chloroplast DNA

Total DNA was isolated as described (Murray and Thompson 1980). For the Southern blots, 3 μg DNA (per lane) was digested with *Sal*I, *Xho*I, and *Sal*I+*Xho*I overnight, resolved on long (20 cm) agarose (0.7%) gels, and transferred to a cationic nylon membrane. The hybridization probes were obtained by PCR of wild-type *Arabidopsis* DNA, followed by random-primer labeling with α-[^{32}P]-dCTP. The *psbA* and *rpoA* probes corresponded to nt 6-1746 and 77335-79091, respectively, of cpDNA (GenBank NC_000932), and were synthesized with oligonucleotides 524 and 525, and 522 and 523, respectively (Table 2). The 18S rDNA probe corresponded to nt 14208429-14210622 of chromosome 3 (GenBank NC_003074), and was amplified with 526 and 527 (Table 3).

For the PCR analysis, 200 ng of genomic DNA was subjected to 25-35 cycles of 94°C (30 sec), 60°C (30 sec), and 70°C (5 min), using Taq polymerase (NEB) and primers 498 and 499 (Table 2), which amplify nt 152241-2786 of cpDNA. The reactions were analyzed on 1% agarose, and the major products were excised for direct sequencing and cloning into *Sma*I-linearized pUC18 or pGC-Blue (Lucigen). The PCR products were also gel-separated into 3 fractions (< 1 kb, 1-3 kb, 3-5 kb) before cloning and sequencing. In parallel, genomic DNA was predigested with I-*Cre*II (Kim *et al*. 2005) for 1 hr at 37 °C to reduce wild-type target DNA, and then used for PCR. The *psbA* primers 510 and 400 (Table 2), which amplify nt 611-947 of cpDNA, were also used for PCR, and the 337-bp product was directly sequenced.

4.12 Analyses of tobacco chloroplast DNA

Total DNA was isolated as described (Murray and Thompson 1980). To verify *MSK57* and *MSK57ix* transplastomic lines, total genomic DNA (1 μg) was digested with *Nco*I overnight, resolved on 0.7% agarose gel, and transferred to a nylon membrane. The hybridization probe was obtained by PCR with the GFP primers 490 and 491 (Table 2), and using pRSET-RIIIG as template. The product (0.7 kb) was purified and subjected to random-primer labeling with α-[^{32}P]-dCTP. The size markers were the 1-kb Plus DNA ladder from Fermentas, which were labeled with γ-[^{32}P]-ATP and polynucleotide kinase (NEB).

4.13 Fluorescence microscopy

Seeds from kanamycin-resistant *Arabidopsis* transformed with pX6RIIIG (*rbcS:I-CreII:gfp*) were germinated on induction medium (containing 10 μM β-estradiol), and grown for 2 weeks under continuous light. Fluorescence microscopy was performed with whole-mounts, or plant sections using a Leica DM IRBE microscope equipped with filters for GFP (excitation at 450-490 nm, and emission at 500-550 nm) and chlorophyll (excitation at 515-560 nm, and detection at >590 nm). The images were captured with a Leica DFC350 FX camera, and processed with Adobe Photoshop.

4.14 Confocal laser scanning microscopy

Subcellular localization of GFP-tagged proteins (rbcS:GFP and aadA:GFP) in tobacco leaf was examined using a Leica SP2 AOBS confocal microscope. GFP was

excited at 488 nm and detected at 508-541 nm; chlorophyll was excited at 633 nm and detected at 654-754 nm. Images were processed with Leica Confocal Software 2.5 and Adobe Photoshop.

4.15 Nucleotide sequence analyses for direct repeats

Whole plastid genomes of *Arabidopsis* and tobacco were analyzed using the software Reputer (Kurtz *et al*. 2001). Repeat sequences around the I-*Cre*II cleavage sites were analyzed using the software Repfinder (Betley *et al*. 2002).

Table 2. Oligonucleotide sequences

No.	Sequence
400	5'-AAGTTTTTCTGATGGTATGCCTCTAGG-3'
485	5'-CTCGAGATGGCTTCCTCTATGCTCTCTTC-3'
486	5'-AGCGCTCCACACCTGCATGCAGTTAAC-3'
488	5'-AGCGCTATGACTACAAAAAAAACAATACAAT-3'
489	5'-CCTAGGAAACGTTATCCAGCCATATGCAG-3'
490	5'-CCTAGGATGGGTAAGGGAGAAGAACTTTTC-3'
491	5'-CCTAGGTACTAGTTTATTTGTATAGTTCATCC-3'
498	5'-CCATGTACGAGGATCCCCAC-3'
499	5'-CGGCAATCCTAGGGTTGCTC-3'
510	5'-CCAAGCCGCTAAGAAGAAATGTAAAG-3'
522	5'-CCTGTAACGTAGGCTTAGCGG-3'
523	5'-AAAGGCCCTGGACTCGGAAG-3'
524	5'-CGAACGACGGGAATTGAACCC-3'
525	5'-TAGTTCCGGGTTCGAGTCCC-3'
526	5'-ATCGGCGCTTGTTCACCTCTC-3'
527	5'-AGCAACGCGCACGACAAGAC-3'
535	5'-GTCTGTGACAATGGAACTGGAA-3'
536	5'-CTTTCTGACCCATACCAACCAT-3'
555	5'-AGGGATATCTATCTAATCCGATC-3'
615	5'-GTATCCATGCGCTTCATATTCG-3'
658	5'-CTAGTACCAAAGTAACCATGAGCAGCTACAATGTTGT-3'
659	5'-CTAGACAACATTGTAGCTGCTCATGGTTACTTTGGTA-3'
672	5'-GGATATCTGAGCCCGCATAGAGCCTTTGATT-3'
673	5'-CCGAATTCGAACAACAATCGGGTG-3'
693	5'-GCTCCTAATGGGACAACATCTG-3'
694	5'-TCAGATACTCTATACGAAACAGG-3'
696	5'-CGTGCATTGATGGGTGCTGAG-3'

References

Abdallah F, Salamini F, and Leister D. (2000) A prediction of the size and evolutionary origin of the proteome of chloroplasts of *Arabidopsis*. *Trends in Plant Science* **5**: 141-142.

Ahlert D, Ruf S, and Bock R. (2003) Plastid protein synthesis is required for plant development in tobacco. *Proceedings of the National Academy of Sciences USA* **100**: 15730-15735.

Ahnesorg P, Smith P, and Jackson SP. (2006) XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. *Cell* **124**: 301-313.

Allen J, and Raven J. (1996) Free-radical-induced mutation vs redox regulation: Costs and benefits of genes in organelles. *Journal of Molecular Evolution* **42**: 482-492.

Allers T, and Lichten M. (2001) Differential timing and control of noncrossover and crossover recombination during meiosis. *Cell* **106**: 47-57.

Allison LA, Simon LD, and Maliga P. (1996) Deletion of *rpoB* reveals a second distinct transcription system in plastids of higher plants. *The EMBO Journal* **15**: 2802-2809.

Allsopp A. (1969) Phylogenetic relationships of the procaryota and the origin of the eucaryotic cell. *New Phytologist* **68**: 591-612.

Anderson DG, and Kowalczykowski SC. (1997) The recombination hot spot χ is a regulatory element that switches the polarity of DNA degradation by the RecBCD enzyme. *Genes & Development* **11**: 571-581.

Anderson DG, Churchill JJ, and Kowalczykowski SC. (1999) A single mutation, RecB(D1080A), eliminates RecA protein loading but not Chi recognition by RecBCD enzyme. *Journal of Biological Chemistry* **274**: 27139 -27144.

Anderson DG, and Kowalczykowski SC. (1997) The translocating RecBCD enzyme stimulates recombination by directing RecA protein onto ssDNA in a χ-regulated manner. *Cell* **90**: 77-86.

Anderson JM. (1999) Insights into the consequences of grana stacking of thylakoid membranes in vascular plants: a personal perspective. *Functional Plant Biology* **26**: 625-639.

Andersson B, and Anderson JM. (1980) Lateral heterogeneity in the distribution of chlorophyll-protein complexes of the thylakoid membranes of spinach chloroplasts. *Biochimica et Biophysica Acta* **593**: 427-440.

Archer EK, and Bonnett HT. (1987) Characterization of a virescent chloroplast mutant of tobacco. *Plant Physiology* **83**: 920-925.

Archibald JM. (2009) The puzzle of plastid evolution. *Current Biology* **19**: R81-R88.

Arnold DA, and Kowalczykowski SC. (2000) Facilitated loading of RecA protein is essential to recombination by RecBCD enzyme. *Journal of Biological Chemistry* **275**: 12261 -12265.

Asai T, Bates DB, and Kogoma T. (1994) DNA replication triggered by double-stranded breaks in *E. coli*: Dependence on homologous recombination functions. *Cell* **78**: 1051-1061.

Ayliffe MA, and Timmis JN. (1992) Plastid DNA sequence homologies in the tobacco nuclear genome. *Molecular and General Genetics* **236**: 105-112.

Bacman SR, Williams SL, and Moraes CT. (2009) Intra- and inter-molecular recombination of mitochondrial DNA after *in vivo* induction of multiple double-strand breaks. *Nucleic Acids Research* **37**: 4218 -4226.

Baena-Gonzalez E, Allahverdiyeva Y, Svab Z, Maliga P, Josse E, Kuntz M, Maenpaa P, and Aro E. (2003) Deletion of the tobacco plastid *psbA* gene triggers an upregulation of the thylakoid-associated NAD(P)H dehydrogenase complex and the plastid terminal oxidase (PTOX). *The Plant Journal* **35**: 704-716.

Bai Y, and Symington LS. (1996) A Rad52 homolog is required for RAD51-independent mitotic recombination in *Saccharomyces cerevisiae*. *Genes & Development* **10**: 2025-2037.

Basile G, Aker M, and Mortimer RK. (1992) Nucleotide sequence and transcriptional regulation of the yeast recombinational repair gene RAD51. *Molecular and Cellular Biology* **12**: 3235-3246.

Bassham JA, Benson AA, and Calvin M. (1950) The path of carbon in photosynthesis. *Journal of Biological Chemistry* **185**: 781 -787.

Belfort M, Derbyshire V, Parker MM, Cousineau B, and Lambowitz AM. (2002) Mobile introns: pathways and proteins. In *Mobile DNA II*, pp. 761-783, ASM Press, Washington.

Bendich AJ. (2004) Circular chloroplast chromosomes: the grand illusion. *The Plant Cell* **16**: 1661-1666.

Bennardo N, Cheng A, Huang N, and Stark JM. (2008) Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair ed. J.E. Haber. *PLoS Genetics* **4**: e1000110.

Bensen RJ, and Warner HR. (1987) Partial purification and characterization of uracil-DNA glycosylase activity from chloroplasts of *Zea mays* seedlings. *Plant Physiology* **84**: 1102-1106.

Bentchikou E, Servant P, Coste G, and Sommer S. (2010) A major role of the RecFOR pathway in DNA double-strand-break repair through ESDSA in *Deinococcus radiodurans*. *PLoS Genetics* **6**: e1000774.

Betley J, Frith MC, Graber JH, Choo S, and Deshler JO. (2002) A ubiquitous and conserved signal for RNA localization in chordates. *Current biology* **12**: 1756-1761.

Beukers R, and Berends W. (1960) Isolation and identification of the irradiation product of thymine. *Biochimica et Biophysica Acta* **41**: 550-551.

Blanck S, Kobbe D, Hartung F, Fengler K, Focke M, and Puchta H. (2009) A SRS2 homolog from *Arabidopsis thaliana* disrupts recombinogenic DNA intermediates and facilitates single strand annealing. *Nucleic Acids Research* **37**: 7163-7176.

Bleuyard J, Gallego ME, and White CI. (2006) Recent advances in understanding of the DNA double-strand break repair machinery of plants. *DNA Repair* **5**: 1-12.

Bock R. (2007) Structure, function, and inheritance of plastid genomes. *Cell and Molecular Biology of Plastids* **19**: 29-63.

Bowater R, and Doherty AJ. (2006) Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining. *PLoS Genetics* **2**: e8.

Boynton J, Gillham N, Harris E, Hosler J, Johnson A, Jones A, Randolph-Anderson B, Robertson D, Klein T, and Shark K. (1988) Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles. *Science* **240**: 1534-1538.

Britt AB. (1996) DNA damage and repair in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**: 75-100.

Brouard J, Otis C, Lemieux C, and Turmel M. (2008) Chloroplast DNA sequence of the green alga *Oedogonium cardiacum* (Chlorophyceae): Unique genome architecture, derived characters shared with the Chaetophorales and novel genes acquired through horizontal transfer. *BMC Genomics* **9**: 290.

Cannon GC, Hedrick LA, and Heinhorst S. (1995) Repair mechanisms of UV-induced DNA damage in soybean chloroplasts. *Plant Molecular Biology* **29**: 1267-1277.

Cao J, Combs C, and Jagendorf AT. (1997) The chloroplast-located homolog of bacterial DNA recombinase. *Plant and Cell Physiology* **38**: 1319 -1325.

Caplan JL, Mamillapalli P, Burch-Smith TM, Czymmek K, and Dinesh-Kumar SP. (2008) Chloroplastic protein NRIP1 mediates innate immune receptor recognition of a viral effector. *Cell* **132**: 449-462.

Cavalier-Smith T. (2002) Chloroplast evolution: secondary symbiogenesis and multiple losses. *Current Biology* **12**: R62-R64.

Cavalier-Smith T. (1982) The origins of plastids. *Biological Journal of the Linnean Society* **17**: 289-306.

Cavalier-Smith T. (2000) Membrane heredity and early chloroplast evolution. *Trends in Plant Science* **5**: 174-182.

Cerutti H, Johnson A, Boynton J, and Gillham N. (1995) Inhibition of chloroplast DNA recombination and repair by dominant negative mutants of *Escherichia coli* RecA. *Molecular and Cellular Biology* **15**: 3003-3011.

Cerutti H, Osman M, Grandoni P, and Jagendorf AT. (1992) A homolog of *Escherichia coli* RecA protein in plastids of higher plants. *Proceedings of the National Academy of Sciences USA* **89**: 8068-8072.

Cerutti H, and Jagendorf AT. (1993) DNA strand-transfer activity in pea (*Pisum sativum* L.) chloroplasts. *Plant Physiology* **102**: 145-153.

Chappell C, Hanakahi LA, Karimi-Busheri F, Weinfeld M, and West SC. (2002) Involvement of human polynucleotide kinase in double-strand break repair by non-homologous end joining. *The EMBO Journal* **21**: 2827-2832.

Chen L, Trujillo K, Ramos W, Sung P, and Tomkinson AE. (2001) Promotion of Dnl4-Catalyzed DNA end-joining by the Rad50/Mre11/Xrs2 and Hdf1/Hdf2 complexes. *Molecular Cell* **8**: 1105-1115.

Chi-Ham CL, Keaton MA, Cannon GC, and Heinhorst S. (2002) The DNA-compacting protein DCP68 from soybean chloroplasts is ferredoxin:sulfite reductase and co-localizes with the organellar nucleoid. *Plant Molecular Biology* **49**: 621-631.

Chomczynski P, and Mackey K. (1995) Modification of the TRI reagent procedure for isolation of RNA from polysaccharide- and proteoglycan-rich sources. *BioTechniques* **19**: 942-945.

Claussen CA, and Long EC. (1999) Nucleic acid recognition by metal complexes of bleomycin. *Chemical Reviews* **99**: 2797-2816.

Clegg MT, Gaut BS, Learn GH, and Morton BR. (1994) Rates and patterns of chloroplast DNA evolution. *Proceedings of the National Academy of Sciences USA* **91**: 6795-6801.

Clough SJ, and Bent AF. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* **16**: 735-743.

Corina LE, Qiu W, Desai A, and Herrin DL. (2009) Biochemical and mutagenic analysis of I-CreII reveals distinct but important roles for both the H-N-H and GIY-YIG motifs. *Nucleic Acids Research* **37**: 5810-5821.

Corneille S, Lutz K, Svab Z, and Maliga P. (2001) Efficient elimination of selectable marker genes from the plastid genome by the CRE-*lox* site-specific recombination system. *The Plant Journal* **27**: 171-178.

Corneille S, Lutz KA, Azhagiri AK, and Maliga P. (2003) Identification of functional *lox* sites in the plastid genome. *The Plant Journal* **35**: 753-762.

Cosner ME, Jansen RK, Palmer JD, and Downie SR. (1997) The highly rearranged chloroplast genome of *Trachelium caeruleum* (Campanulaceae): multiple inversions, inverted repeat expansion and contraction, transposition, insertions/deletions, and several repeat families. *Current Genetics* **31**: 419-429.

Cox MM, Keck JL, and Battista JR. (2010) Rising from the ashes: DNA repair in *Deinococcus radiodurans*. *PLoS Genetics* **6**: e1000815.

Cromie GA, Connelly JC, and Leach DR. (2001) Recombination at double-strand breaks and DNA ends: Conserved mechanisms from phage to humans. *Molecular Cell* **8**: 1163-1174.

Cséplő Á, and Maliga P. (1984) Large scale isolation of maternally inherited lincomycin resistance mutations, in diploid *Nicotiana plumbaginifolia* protoplast cultures. *Molecular and General Genetics* **196**: 407-412.

Cunningham FX, and Gantt E. (1998) Genes and enzymes of carotenoid biosynthesis in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **49**: 557-583.

Davies AA, Friedberg EC, Tomkinson AE, Wood RD, and West SC. (1995) Role of the Rad1 and Rad10 proteins in nucleotide excision repair and recombination. *Journal of Biological Chemistry* **270**: 24638 -24641.

Davis AP, and Symington LS. (2004) Rad51-dependent break-induced replication in yeast. *Molecular and Cellular Biology* **24**: 2344-2351.

Decottignies A. (2007) Microhomology-mediated end joining in fission yeast is repressed by Pku70 and relies on genes involved in homologous recombination. *Genetics* **176**: 1403-1415.

Della M. (2004) Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine. *Science* **306**: 683-685.

Deng XW, Wing RA, and Gruissem W. (1989) The chloroplast genome exists in multimeric forms. *Proceedings of the National Academy of Sciences USA* **86**: 4156-4160.

Dixon DA, and Kowalczykowski SC. (1993) The recombination hotspot χ is a regulatory sequence that acts by attenuating the nuclease activity of the *E. coli* RecBCD enzyme. *Cell* **73**: 87-96.

Doherty AJ, Jackson SP, and Weller GR. (2001) Identification of bacterial homologues of the Ku DNA repair proteins. *FEBS Letters* **500**: 186-188.

Drescher A, Ruf S, Calsa TJ, Carrer H, and Bock R. (2000) The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. *The Plant Journal* **22**: 97-104.

Dufourmantel N, Dubald M, Matringe M, Canard H, Garcon F, Job C, Kay E, Wisniewski J, Ferullo J, Pelissier B, Sailland A, and Tissot G. (2007) Generation

and characterization of soybean and marker-free tobacco plastid transformants over-expressing a bacterial 4-hydroxyphenylpyruvate dioxygenase which provides strong herbicide tolerance. *Plant Biotechnology Journal* **5**: 118-133.

Dufourmantel N, Pelissier B, Garcon F, Peltier G, Ferullo J, and Tissot G. (2004) Generation of fertile transplastomic soybean. *Plant Molecular Biology* **55**: 479-489.

Dürrenberger F, Thompson AJ, Herrin DL, and Rochaix JD. (1996) Double strand break-induced recombination in *Chlamydomonas reinhardtii* chloroplasts. *Nucleic Acids Research* **24**: 3323-3331.

Eggleston AK, and West SC. (1997) Recombination initiation: easy as A, B, C, D... χ? *Current Biology* **7**: R745-R749.

Eisen JA, and Hanawalt PC. (1999) A phylogenomic study of DNA repair genes, proteins, and processes. *DNA Repair* **435**: 171-213.

Elias BA, and Givan CV. (1977) Alpha-ketoglutarate supply for amino acid synthesis in higher plant chloroplasts: intrachloroplastic localization of NADP-specific isocitrate dehydrogenase. *Plant Physiology* **59**: 738-740.

Fishman-Lobell J, and Haber J. (1992) Removal of nonhomologous DNA ends in double-strand break recombination: the role of the yeast ultraviolet repair gene RAD1. *Science* **258**: 480-484.

Flott S, Alabert C, Toh GW, Toth R, Sugawara N, Campbell DG, Haber JE, Pasero P, and Rouse J. (2007) Phosphorylation of Slx4 by Mec1 and Tel1 regulates the single-strand annealing mode of DNA repair in budding yeast. *Molecular and Cellular Biology* **27**: 6433-6445.

Franklin S, Ngo B, Efuet E, and Mayfield SP. (2002) Development of a GFP reporter gene for *Chlamydomonas reinhardtii* chloroplast. *The Plant Journal* **30**: 733-744.

Friesen LF, Nelson AG, and Van Acker RC. (2003) Evidence of contamination of pedigreed canola (*Brassica napus*) seedlots in western canada with genetically engineered herbicide resistance traits. *Agronomy Journal* **95**: 1342-1347.

Glover L, McCulloch R, and Horn D. (2008) Sequence homology and microhomology dominate chromosomal double-strand break repair in African trypanosomes. *Nucleic Acids Research* **36**: 2608 -2618.

Gottlieb TM, and Jackson SP. (1993) The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen. *Cell* **72**: 131-142.

Goulding S, Wolfe K, Olmstead R, and Morden C. (1996) Ebb and flow of the chloroplast inverted repeat. *Molecular and General Genetics* **252**: 195-206.

Grawunder U, Wilm M, Wu X, Kulesza P, Wilson TE, Mann M, and Lieber MR. (1997) Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature* **388**: 492-495.

Gray MW, and Doolittle WF. (1982) Has the endosymbiont hypothesis been proven? *Microbiological Reviews* **46**: 1-42.

Gruen M, Chang K, Serbanescu I, and Liu DR. (2002) An *in vivo* selection system for homing endonuclease activity. *Nucleic Acids Research* **30**: e29.

GuhaMajumdar M, and Sears BB. (2005) Chloroplast DNA base substitutions: an experimental assessment. *Molecular Genetics and Genomics* **273**: 177-183.

Gutman BL, and Niyogi KK. (2009) Evidence for base excision repair of oxidative DNA damage in chloroplasts of *Arabidopsis thaliana*. *Journal of Biological Chemistry* **284**: 17006 -17012.

Haber JE. (2008) Alternative endings. *Proceedings of the National Academy of Sciences USA* **105**: 405 -406.

Hajdukiewicz PTJ, Gilbertson L, and Staub JM. (2001) Multiple pathways for Cre/*lox*-mediated recombination in plastids. *The Plant Journal* **27**: 161-170.

Hallick RB, Hong L, Drager RG, Favreau MR, Monfort A, Orsat B, Spielmann A, and Stutz E. (1993) Complete sequence of *Euglena gracilis* chloroplast DNA. *Nucleic Acids Research* **21**: 3537-3544.

Halliwell B, and Aruoma OI. (1991) DNA damage by oxygen-derived species Its mechanism and measurement in mammalian systems. *FEBS Letters* **281**: 9-19.

Hanson MR, and Sattarzadeh A. (2008) Dynamic morphology of plastids and stromules in angiosperm plants. *Plant, Cell & Environment* **31**: 646-657.

Haslberger AG. (2000) Genetic technologies:monitoring and labeling for genetically modified products. *Science* **287**: 431-432.

Heacock M, Spangler E, Riha K, Puizina J, and Shippen DE. (2004) Molecular analysis of telomere fusions in *Arabidopsis*: multiple pathways for chromosome end-joining. *The EMBO Journal* **23**: 2304-2313.

Heinhorst S, Chi-Ham CL, Adamson SW, and Cannon GC. (2004) The somatic inheritance of plant organelles. In *Molecular Biology and Biotechnology of Plant Organelles*, pp. 37-92, Springer Netherlands.

Hell R. (1997) Molecular physiology of plant sulfur metabolism. *Planta* **202**: 138-148.

Herrmann G, Lindahl T, and Schär P. (1998) *Saccharomyces cerevisiae* LIF1: a function involved in DNA double-strand break repair related to mammalian XRCC4. *The EMBO Journal* **17**: 4188-4198.

Honda M, Fujisawa T, Shibata T, and Mikawa T. (2008) RecR forms a ring-like tetramer that encircles dsDNA by forming a complex with RecF. *Nucleic Acids Research* **36**: 5013-5020.

Hopfner K, Craig L, Moncalian G, Zinkel RA, Usui T, Owen BAL, Karcher A, Henderson B, Bodmer J, McMurray CT, Carney JP, Petrini JHJ, and Tainer JA. (2002) The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. *Nature* **418**: 562-566.

Huang, Klaus, Herz, Zou, Koop, and Golds. (2002) Efficient plastid transformation in tobacco using the *aphA-6* gene and kanamycin selection. *Molecular Genetics and Genomics* **268**: 19-27.

Iamtham S, and Day A. (2000) Removal of antibiotic resistance genes from transgenic tobacco plastids. *Nature Biotechnology* **18**: 1172-1176.

Ira G, Malkova A, Liberi G, Foiani M, and Haber JE. (2003) Srs2 and Sgs1-Top3 suppress crossovers during double-strand break repair in yeast. *Cell* **115**: 401-411.

Ivancic-Bace I, Peharec P, Moslavac S, Skrobot N, Salaj-Smic E, and Brcic-Kostic K. (2003) RecFOR function is required for DNA repair and recombination in a RecA loading-deficient *recB* mutant of *Escherichia coli*. *Genetics* **163**: 485-494.

Ivancic-Bace I, Salaj-Smic E, and Brcic-Kostic K. (2005) Effects of *recJ*, *recQ*, and *recFOR* mutations on recombination in nuclease-deficient *recB recD* double mutants of *Escherichia coli*. *The Journal of Bacteriology* **187**: 1350-1356.

Iwasaki H, Takahagi M, Nakata A, and Shinagawa H. (1992) *Escherichia coli* RuvA and RuvB proteins specifically interact with Holliday junctions and promote branch migration. *Genes & Development* **6**: 2214-2220.

Jan van Wijk K. (2000) Proteomics of the chloroplast: experimentation and prediction. *Trends in Plant Science* **5**: 420-425.

Jansen R, Kaittanis C, Saski C, Lee S, Tomkins J, Alverson A, and Daniell H. (2006) Phylogenetic analyses of *Vitis* (Vitaceae) based on complete chloroplast genome sequences: effects of taxon sampling and phylogenetic methods on resolving relationships among rosids. *BMC Evolutionary Biology* **6**: 32.

Jeong SY, Rose A, and Meier I. (2003) MFP1 is a thylakoid-associated, nucleoid-binding protein with a coiled-coil structure. *Nucleic Acids Research* **31**: 5175-5185.

Kanevski I, and Maliga P. (1994) Relocation of the plastid *rbcL* gene to the nucleus yields functional ribulose-1,5-bisphosphate carboxylase in tobacco chloroplasts. *Proceedings of the National Academy of Sciences USA* **91**: 1969-1973.

Kegel A, Sjöstrand JO, and Aström SU. (2001) Nej1p, a cell type-specific regulator of nonhomologous end joining in yeast. *Current Biology* **11**: 1611-1617.

Khakhlova O, and Bock R. (2006) Elimination of deleterious mutations in plastid genomes by gene conversion. *The Plant Journal* **46**: 85-94.

Khan MS, and Maliga P. (1999) Fluorescent antibiotic resistance marker for tracking plastid transformation in higher plants. *Nature Biotechnology* **17**: 910-915.

Kim H, Corina LE, Suh J, and Herrin DL. (2005) Expression, purification, and biochemical characterization of the intron-encoded endonuclease, I-CreII. *Protein Expression and Purification* **44**: 162-172.

Kittiwongwattana C, Lutz K, Clark M, and Maliga P. (2007) Plastid marker gene excision by the phiC31 phage site-specific recombinase. *Plant Molecular Biology* **64**: 137-143.

Kmiec EB, Johnson C, and May GD. (2001) Chloroplast lysates support directed mutagenesis via modified DNA and chimeric RNA/DNA oligonucleotides. *The Plant Journal* **27**: 267-274.

Kobayashi T, Takahara M, Miyagishima S, Kuroiwa H, Sasaki N, Ohta N, Matsuzaki M, and Kuroiwa T. (2002) Detection and localization of a chloroplast-encoded HU-like protein that organizes chloroplast nucleoids. *The Plant Cell* **14**: 1579-1589.

Kode V, Mudd EA, Iamtham S, and Day A. (2006) Isolation of precise plastid deletion mutants by homology-based excision: a resource for site-directed mutagenesis, multi-gene changes and high-throughput plastid transformation. *The Plant Journal* **46**: 901-909.

Kohl S, and Bock R. (2009) Transposition of a bacterial insertion sequence in chloroplasts. *The Plant Journal* **58**: 423-436.

Kohler RH, Cao J, Zipfel WR, Webb WW, and Hanson MR. (1997) Exchange of protein molecules through connections between higher plant plastids. *Science* **276**: 2039-2042.

Kolodner R, and Tewari K. (1975) The molecular size and conformation of the chloroplast DNA from higher plants. *Biochimica et Biophysica Acta* **402**: 372-390.

Kolodner R, and Tewari KK. (1979) Inverted repeats in chloroplast DNA from higher plants. *Proceedings of the National Academy of Sciences USA* **76**: 41 -45.

Korpelainen H. (2004) The evolutionary processes of mitochondrial and chloroplast genomes differ from those of nuclear genomes. *Naturwissenschaften* **91**: 505-518.

Krebbers E, Seurinck J, Herdies L, Cashmore AR, and Timko MP. (1988) Four genes in two diverged subfamilies encode the ribulose-1,5-bisphosphate carboxylase small subunit polypeptides of *Arabidopsis thaliana*. *Plant Molecular Biology* **11**: 745-759.

Kumar S, Dhingra A, and Daniell H. (2004a) Plastid-expressed betaine aldehyde dehydrogenase gene in carrot cultured cells, roots, and leaves confers enhanced salt tolerance. *Plant Physiology* **136**: 2843-2854.

Kumar S, Dhingra A, and Daniell H. (2004b) Stable transformation of the cotton plastid genome and maternal inheritance of transgenes. *Plant Molecular Biology* **56**: 203-216.

Kuroda H, and Maliga P. (2003) The plastid *clpP1* protease gene is essential for plant development. *Nature* **425**: 86-89.

Kuroiwa T. (1991) The replication, differentiation, and inheritance of plastids with emphasis on the concept of organelle nuclei. *International Review of Cytology* **128**: 1-62.

Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, and Giegerich R. (2001) REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Research* **29**: 4633 -4642.

Kwon T, Huq E, and Herrin DL. (2010) Microhomology-mediated and nonhomologous repair of a double-strand break in the chloroplast genome of *Arabidopsis*. *Proceedings of the National Academy of Sciences USA* **107**: 13954 -13959.

La Claire JW, and Wang J. (2004) Structural characterization of the terminal domains of linear plasmid-like DNA from the green alga *Ernodesmis* (Chlorophyta). *Journal of Phycology* **40**: 1089-1097.

Lam H, Coschigano KT, Oliveira IC, Melo-Oliveira R, and Coruzzi GM. (1996) The molecular-genetics of nitrogen assimilation into amino acids in higher plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**: 569-593.

Lee DW, Lee S, Lee G, Lee KH, Kim S, Cheong G, and Hwang I. (2006) Functional characterization of sequence motifs in the transit peptide of *Arabidopsis* small subunit of rubisco. *Plant Physiology* **140**: 466-483.

Lee SE, Moore JK, Holmes A, Umezu K, Kolodner RD, and Haber JE. (1998) *Saccharomyces* Ku70, Mre11/Rad50, and RPA proteins regulate adaptation to G2/M arrest after DNA damage. *Cell* **94**: 399-409.

Legen J, Wanner G, Herrmann RG, Small I, and Schmitz-Linneweber C. (2007) Plastid tRNA genes *trnC*-GCA and *trnN*-GUU are essential for plant cell development. *The Plant Journal* **51**: 751-762.

Lelivelt CLC, McCabe MS, Newell CA, Desnoo CB, van Dun KMP, Birch-Machin I, Gray JC, Mills KHG, and Nugent JM. (2005) Stable plastid transformation in lettuce (*Lactuca sativa* L.). *Plant Molecular Biology* **58**: 763-774.

Liang L, Deng L, Nguyen SC, Zhao X, Maulion CD, Shao C, and Tischfield JA. (2008) Human DNA ligases I and III, but not ligase IV, are required for microhomology-mediated end joining of DNA double-strand breaks. *Nucleic Acids Research* **36**: 3297-3310.

Lilly JW, Havey MJ, Jackson SA, and Jiang J. (2001) Cytogenomic analyses reveal the structural plasticity of the chloroplast genome in higher plants. *The Plant Cell* **13**: 245-254.

Lin FL, Sperle K, and Sternberg N. (1984) Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process. *Molecular and Cellular Biology* **4**: 1020-1034.

Liu JW, and Rose RJ. (1992) The spinach chloroplast chromosome is bound to the thylakoid membrane in the region of the inverted repeat. *Biochemical and Biophysical Research Communications* **184**: 993-1000.

Liu Y, Masson J, Shah R, O'Regan P, and West SC. (2004) Rad51c is required for Holliday junction processing in mammalian cells. *Science* **303**: 243-246.

Llorente B, and Symington LS. (2004) The Mre11 nuclease is not required for 5' to 3' resection at multiple HO-induced double-strand breaks. *Molecular and Cellular Biology* **24**: 9682-9694.

Lloyd RG. (1991) Conjugational recombination in resolvase-deficient *ruvC* mutants of *Escherichia coli* K-12 depends on *recG*. *The Journal of Bacteriology* **173**: 5414-5418.

Lloyd RG, and Sharples GJ. (1993) Dissociation of synthetic Holliday junctions by *E. coli* RecG protein. *The EMBO Journal* **12**: 17-22.

Logan DC. (2010) Mitochondrial fusion, division and positioning in plants. *Biochemical Society Transactions* **38**: 789-795.

Lutz KA, Bosacchi MH, and Maliga P. (2006) Plastid marker-gene excision by transiently expressed CRE recombinase. *The Plant Journal* **45**: 447-456.

Lutz KA, Azhagiri AK, Tungsuchat-Huang T, and Maliga P. (2007) A guide to choosing vectors for transformation of the plastid genome of higher plants. *Plant Physiology* **145**: 1201-1210.

Lutz KA, Svab Z, and Maliga P. (2006) Construction of marker-free transplastomic tobacco using the Cre-*lox*P site-specific recombination system. *Nature Protocols* **1**: 900-910.

Ma J, Kim EM, Haber JE, and Lee SE. (2003) Yeast Mre11 and Rad1 proteins define a Ku-independent mechanism to repair double-strand breaks lacking overlapping end sequences. *Molecular and Cellular Biology* **23**: 8820-8828.

Maliga P. (2004) Plastid transformation in higher plants. *Annual Review of Plant Biology* **55**: 289-313.

Maréchal A, Parent J, Veronneau-Lafortune F, Joyeux A, Lang BF, and Brisson N. (2009) Whirly proteins maintain plastid genome stability in *Arabidopsis*. *Proceedings of the National Academy of Sciences USA* **106**: 14693-14698.

Martin W, and Herrmann RG. (1998) Gene transfer from organelles to the nucleus: how much, what happens, and why? *Plant Physiology* **118**: 9-17.

Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, and Penny D. (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proceedings of the National Academy of Sciences USA* **99**: 12246 -12251.

Maul JE, Lilly JW, Cui L, dePamphilis CW, Miller W, Harris EH, and Stern DB. (2002) The *Chlamydomonas reinhardtii* plastid chromosome: Islands of genes in a sea of repeats. *The Plant Cell* **14**: 2659-2679.

Mezard C, and Nicolas A. (1994) Homologous, homeologous, and illegitimate repair of double-strand breaks during transformation of a wild-type strain and a rad52 mutant strain of *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **14**: 1278-1292.

Moore J, and Haber J. (1996) Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **16**: 2164-2173.

Moreira D, Le Guyader H, and Philippe H. (2000) The origin of red algae and the evolution of chloroplasts. *Nature* **405**: 69-72.

Murashige T, and Skoog F. (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia Plantarum* **15**: 473-497.

Murray M, and Thompson W. (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* **8**: 4321-4326.

Mustardy L, Buttle K, Steinbach G, and Garab G. (2008) The three-dimensional network of the thylakoid membranes in plants: quasihelical model of the granum-stroma assembly. *The Plant Cell* **20**: 2552-2557.

Naciri Y, and Manen J. (2010) Potential DNA transfer from the chloroplast to the nucleus in *Eryngium alpinum*. *Molecular Ecology Resources* **10**: 728-731.

Nakamura Y, Gojobori T, and Ikemura T. (2000) Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Research* **28**: 292.

Niewolik D, Pannicke U, Lu H, Ma Y, Wang LV, Kulesza P, Zandi E, Lieber MR, and Schwarz K. (2006) DNA-PKcs dependence of Artemis endonucleolytic activity, differences between hairpins and 5′ or 3′ overhangs. *Journal of Biological Chemistry* **281**: 33900 -33909.

Nimonkar AV, Özsoy AZ, Genschel J, Modrich P, and Kowalczykowski SC. (2008) Human exonuclease 1 and BLM helicase interact to resect DNA and initiate DNA repair. *Proceedings of the National Academy of Sciences USA* **105**: 16906 -16911.

Noskov VN, Segall-Shapiro TH, and Chuang R. (2010) Tandem repeat coupled with endonuclease cleavage (TREC): a seamless modification tool for genome engineering in yeast. *Nucleic Acids Research* **38**: 2570 -2576.

Noutsos C, Richly E, and Leister D. (2005) Generation and evolutionary fate of insertions of organelle DNA in the nuclear genomes of flowering plants. *Genome Research* **15**: 616 -628.

Odom OW, Shenkenberg DL, Garcia JA, and Herrin DL. (2004) A horizontally acquired group II intron in the chloroplast *psbA* gene of a psychrophilic Chlamydomonas: *in vitro* self-splicing and genetic evidence for maturase activity. *RNA* **10**: 1097-1107.

Odom OW, Baek K, Dani RN, and Herrin DL. (2008) *Chlamydomonas* chloroplasts can use short dispersed repeats and multiple pathways to repair a double-strand break in the genome. *The Plant Journal* **53**: 842-853.

Odom OW, Holloway SP, Deshpande NN, Lee J, and Herrin DL. (2001) Mobile self-splicing Group I introns from the *psbA* gene of *Chlamydomonas reinhardtii*: highly efficient homing of an exogenous intron containing its own promoter. *Molecular and Cellular Biology* **21**: 3472-3481.

Ow DW. (2002) Recombinase-directed plant transformation for the post-genomic era. *Plant Molecular Biology* **48**: 183-200.

Palmer JD. (1983) Chloroplast DNA exists in two orientations. *Nature* **301**: 92-93.

Paques F, and Haber JE. (1999) Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiology and Molecular Biology Reviews* **63**: 349-404.

Pardo B, Gómez-González B, and Aguilera A. (2009) DNA repair in mammalian cells. *Cellular and Molecular Life Sciences* **66**: 1039-1056.

Perez C, Guyot V, Cabaniols J, Gouble A, Micheaux B, Smith J, Leduc S, Pâques F, and Duchateau P. (2005) Factors affecting double-strand break-induced homologous recombination in mammalian cells. *BioTechniques* **39**: 109-115.

Pfitzinger H, Weil JH, Pillay DTN, and Guillemaut P. (1990) Codon recognition mechanisms in plant chloroplasts. *Plant Molecular Biology* **14**: 805-814.

Puchta H. (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *Journal of Experimental Botany* **56**: 1-14.

Ragu S, Faye G, Iraqui I, Masurel-Heneman A, Kolodner RD, and Huang M. (2007) Oxygen metabolism and reactive oxygen species cause chromosomal rearrangements and cell death. *Proceedings of the National Academy of Sciences USA* **104**: 9747 -9752.

Robert T, Dervins D, Fabre F, and Gangloff S. (2006) Mrc1 and Srs2 are major actors in the regulation of spontaneous crossover. *The EMBO Journal* **25**: 2837-2846.

Robertson EJ, Pyke KA, and Leech RM. (1995) arc6, an extreme chloroplast division mutant of *Arabidopsis* also alters proplastid proliferation and morphology in shoot and root apices. *Journal of Cell Science* **108**: 2937-2944.

Rocha EPC, Cornet E, and Michel B. (2005) Comparative and evolutionary analysis of the bacterial homologous recombination systems. *PLoS Genetics* **1**.

Rochaix J. (1997) Chloroplast reverse genetics: new insights into the function of plastid genes. *Trends in Plant Science* **2**: 419-425.

Roff R. (2009) No alternative? The politics and history of non-GMO certification. *Agriculture and Human Values* **26**: 351-363.

Rogers SG, Horsch RB, and Fraley RT. (1986) Gene transfer in plants: production of transformed plants using Ti plasmid vectors. *Methods in Enzymology* **118**: 627-640.

Rose RJ. (1979) The association of chloroplast DNA with photosynthetic membrane vesicles from spinach chloroplasts. *Journal of Cell Science* **36**: 169-183.

Roughan PG, Holland R, and Slack CR. (1979) On the control of long-chain-fatty acid synthesis in isolated intact spinach (*Spinacia oleracea*) chloroplasts. *Biochemical Journal* **184**: 193-202.

Rowan BA, Oldenburg DJ, and Bendich AJ. (2004) The demise of chloroplast DNA in *Arabidopsis*. *Current Genetics* **46**: 176-181.

Ruf S, Hermann M, Berger IJ, Carrer H, and Bock R. (2001) Stable genetic transformation of tomato plastids and expression of a foreign protein in fruit. *Nature Biotechnology* **19**: 870-875.

Ruf S, Karcher D, and Bock R. (2007) Determining the transgene containment level provided by chloroplast transformation. *Proceedings of the National Academy of Sciences USA* **104**: 6998 -7002.

Sakai A, Takano H, and Kuroiwa T. (2004) Organelle nuclei in higher plants: structure, composition, function, and evolution. *International Review of Cytology* **238**: 59-118.

Sanchez H, and Alonso JC. (2005) Bacillus subtilis RecN binds and protects 3′-single-stranded DNA extensions in the presence of ATP. *Nucleic Acids Research* **33**: 2343 -2350.

Saparbaev M, Prakash L, and Prakash S. (1996) Requirement of mismatch repair genes MSH2 and MSH3 in the RAD1-RAD10 pathway of mitotic recombination in *Saccharomyces cerevisiae*. *Genetics* **142**: 727-736.

Sargent R, Brenneman M, and Wilson J. (1997) Repair of site-specific double-strand breaks in a mammalian chromosome by homologous and illegitimate recombination. *Molecular and Cellular Biology* **17**: 267-277.

Sasaki Y, Hakamada K, Suama Y, Nagano Y, Furusawa I, and Matsuno R. (1993) Chloroplast-encoded protein as a subunit of acetyl-CoA carboxylase in pea plant. *Journal of Biological Chemistry* **268**: 25118 -25123.

Saski C, Lee S, Daniell H, Wood T, Tomkins J, Kim H, and Jansen R. (2005) Complete chloroplast genome sequence of *Glycine max* and comparative analyses with other legume genomes. *Plant Molecular Biology* **59**: 309-322.

Sato N, Albrieux C, Joyard J, Douce R, and Kuroiwa T. (1993) Detection and characterization of a plastid envelope DNA-binding protein which may anchor plastid nucleoids. *The EMBO Journal* **12**: 555-561.

Sato N, Misumi O, Shinada Y, Sasaki M, and Yoine M. (1997) Dynamics of localization and protein composition of plastid nucleoids in light-grown pea seedlings. *Protoplasma* **200**: 163-173.

Sato N, Ohshima K, Watanabe A, Ohta N, Nishiyama Y, Joyard J, and Douce R. (1998) Molecular characterization of the PEND protein, a novel bZIP protein present in the envelope membrane that is the site of nucleoid replication in developing plastids. *The Plant Cell* **10**: 859-872.

Sato S, Nakamura Y, Kaneko T, Asamizu E, and Tabata S. (1999) Complete structure of the chloroplast genome of *Arabidopsis thaliana*. *DNA Research* **6**: 283 -290.

Schild-Poulter C, Pope L, Giffin W, Kochan JC, Ngsee JK, Traykova-Andonova M, and Haché RJ. (2001) The binding of Ku antigen to homeodomain proteins promotes their phosphorylation by DNA-dependent protein kinase. *The Journal of Biological Chemistry* **276**: 16848-16856.

Scholes G, Ward J, and Weiss J. (1960) Mechanism of the radiation-induced degradation of nucleic acids. *Journal of Molecular Biology* **2**: 379-391.

Scott NS, and Possingham JV. (1980) Chloroplast DNA in expanding spinach leaves. *Journal of Experimental Botany* **31**: 1081 -1092.

Selby CP, and Sancar A. (2006) A cryptochrome/photolyase class of enzymes with single-stranded DNA-specific photolyase activity. *Proceedings of the National Academy of Sciences USA* **103**: 17696 -17700.

Shah R, Bennett RJ, and West SC. (1994) Activation of RuvC Holliday junction resolvase *in vitro*. *Nucleic Acids Research* **22**: 2490-2497.

Shahmuradov IA, Akbarova YY, Solovyev VV, and Aliyev JA. (2003) Abundance of plastid DNA insertions in nuclear genomes of rice and *Arabidopsis*. *Plant Molecular Biology* **52**: 923-934.

Shapiro JA, and von Sternberg R. (2005) Why repetitive DNA is essential to genome function. *Biological Reviews* **80**: 227-250.

Shimda H, and Sugiuro M. (1991) Fine structural features of the chloroplast genome: comparison of the sequenced chloroplast genomes. *Nucleic Acids Research* **19**: 983 -995.

Sibanda BL, Critchlow SE, Begun J, Pei XY, Jackson SP, Blundell TL, and Pellegrini L. (2001) Crystal structure of an Xrcc4-DNA ligase IV complex. *Nature Structural Biology* **8**: 1015-1019.

Sidorov VA, Kasten D, Pang S, Hajdukiewicz PTJ, Staub JM, and Nehra N. (1999) Stable chloroplast transformation in potato: use of green fluorescent protein as a plastid marker. *The Plant Journal* **19**: 209-216.

Siebert R, and Puchta H. (2002) Efficient repair of genomic double-strand breaks by homologous recombination between directly repeated sequences in the plant genome. *The Plant Cell* **14**: 1121-1131.

Sikdar SR, Serino G, Chaudhuri S, and Maliga P. (1998) Plastid transformation in *Arabidopsis thaliana*. *Plant Cell Reports* **18**: 20-24.

Skarjinskaia M, Svab Z, and Maliga P. (2003) Plastid transformation in *Lesquerella fendleri*, an oilseed Brassicacea. *Transgenic Research* **12**: 115-122.

Small GD, and Greimann CS. (1977) Photoreactivation and dark repair of ultraviolet light-induced pyrimidine dimers in chloroplast DNA. *Nucleic Acids Research* **4**: 2893-2902.

Song B, and Sung P. (2000) Functional interactions among yeast Rad51 recombinase, Rad52 mediator, and Replication Protein A in DNA strand exchange. *Journal of Biological Chemistry* **275**: 15895 -15904.

Stanga JP, Boonsirichai K, Sedbrook JC, Otegui MS, and Masson PH. (2009) A role for the TOC complex in *Arabidopsis* root gravitropism. *Plant Physiology* **149**: 1896-1905.

Stasolla C, Katahira R, Thorpe TA, and Ashihara H. (2003) Purine and pyrimidine nucleotide metabolism in higher plants. *Journal of Plant Physiology* **160**: 1271-1295.

Staub JM, and Maliga P. (1994) Extrachromosomal elements in tobacco plastids. *Proceedings of the National Academy of Sciences USA* **91**: 7468-7472.

Stephanou NC. (2007) Mycobacterial nonhomologous end joining mediates mutagenic repair of chromosomal double-strand DNA breaks. *Journal of Bacteriology* **189**: 5237-5246.

Stitt M, and Heldt H. (1981) Simultaneous synthesis and degradation of starch in spinach chloroplasts in the light. *Biochimica et Biophysica Acta* **638**: 1-11.

Sugawara N, Ira G, and Haber JE. (2000) DNA length dependence of the single-strand annealing pathway and the role of *Saccharomyces cerevisiae* Rad59 in double-strand break repair. *Molecular and Cellular Biology* **20**: 5300-5309.

Sugiura M. (1992) The chloroplast genome. *Plant Molecular Biology* **19**: 149-168.

Sung P. (1994) Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. *Science* **265**: 1241-1243.

Suss KH, Arkona C, Manteuffel R, and Adler K. (1993) Calvin cycle multienzyme complexes are bound to chloroplast thylakoid membranes of higher plants in situ. *Proceedings of the National Academy of Sciences USA* **90**: 5514 -5518.

Svab Z, Hajdukiewicz P, and Maliga P. (1990) Stable transformation of plastids in higher plants. *Proceedings of the National Academy of Sciences USA* **87**: 8526 -8530.

Svab Z, and Maliga P. (2007) Exceptional transmission of plastids and mitochondria from the transplastomic pollen parent and its impact on transgene containment. *Proceedings of the National Academy of Sciences USA* **104**: 7003 -7008.

Swarbreck D, Wilks C, Lamesch P, Berardini TZ, Garcia-Hernandez M, Foerster H, Li D, Meyer T, Muller R, Ploetz L, Radenbaugh A, Singh S, Swing V, Tissier C, Zhang P, and Huala E. (2008) The *Arabidopsis* Information Resource (TAIR): gene structure and function annotation. *Nucleic Acids Research* **36**: D1009 - D1014.

Szostak JW, Orr-Weaver TL, Rothstein RJ, and Stahl FW. (1983) The double-strand-break repair model for recombination. *Cell* **33**: 25-35.

Tamura K, Adachi Y, Chiba K, Oguchi K, and Takahashi H. (2002) Identification of Ku70 and Ku80 homologues in *Arabidopsis thaliana*: evidence for a role in the repair of DNA double-strand breaks. *The Plant Journal* **29**: 771-781.

Thomas MR, and Rose RJ. (1983) Plastid number and plastid structural changes associated with tobacco mesophyll protoplast culture and plant regeneration. *Planta* **158**: 329-338.

Thorpe HM, and Smith MC. (1998) *In vitro* site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proceedings of the National Academy of Sciences USA* **95**: 5505-5510.

Thorpe HM, Wilson SE, and Smith MC. (2000) Control of directionality in the site-specific recombination system of the *Streptomyces* phage phiC31. *Molecular Microbiology* **38**: 232-241.

Timmis JN, and Scott NS. (1983) Sequence homology between spinach nuclear and chloroplast genomes. *Nature* **305**: 65-67.

Tominaga M, Kinoshita T, and Shimazaki K. (2001) Guard-cell chloroplasts provide ATP required for H^+ pumping in the plasma membrane and stomatal opening. *Plant and Cell Physiology* **42**: 795 -802.

Trujillo KM, Roh DH, Chen L, Van Komen S, Tomkinson A, and Sung P. (2003) Yeast Xrs2 binds DNA and helps target Rad50 and Mre11 to DNA ends. *Journal of Biological Chemistry* **278**: 48957 -48964.

Tseng H, and Tomkinson AE. (2004) Processing and joining of DNA ends coordinated by interactions among Dnl4/Lif1, Pol4, and FEN-1. *Journal of Biological Chemistry* **279**: 47580 -47588.

Tsubouchi H, and Ogawa H. (2000) Exo1 roles for repair of DNA double-strand breaks and meiotic crossing over in *Saccharomyces cerevisiae*. *Molecular Biology of the Cell* **11**: 2221-2233.

Tsuruya K, Suzuki M, Plader W, Sugita C, and Sugita M. (2006) Chloroplast transformation reveals that tobacco *ycf5* is involved in photosynthesis. *Acta Physiologiae Plantarum* **28**: 365-372.

Umezu K, and Kolodner RD. (1994) Protein interactions in genetic recombination in *Escherichia coli*. Interactions involving RecO and RecR overcome the inhibition of RecA by single-stranded DNA-binding protein. *Journal of Biological Chemistry* **269**: 30005 -30013.

Vahrenholz C, Riemen G, Pratje E, Dujon B, and Michaelis G. (1993) Mitochondrial DNA of *Chlamydomonas reinhardtii*: the structure of the ends of the linear 15.8-kb genome suggests mechanisms for DNA replication. *Current Genetics* **24**: 241-247.

Von Wettstein D, Gough S, and Kannangara C. (1995) Chlorophyll biosynthesis. *The Plant Cell* **7**: 1039-1057.

Warwick SI, Beckie HJ, and Hall LM. (2009) Gene flow, invasiveness, and ecological impact of genetically modified crops. *Annals of the New York Academy of Sciences* **1168**: 72-99.

Weller GR, Kysela B, Roy R, Tonkin LM, Scanlan E, Della M, Devine SK, Day JP, Wilkinson A, di Fagagna FD, Devine KM, Bowater RP, Jeggo PA, Jackson SP, and Doherty AJ. (2002) Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* **297**: 1686-1689.

West CE, Waterworth WM, Story GW, Sunderland PA, Jiang Q, and Bray CM. (2002) Disruption of the *Arabidopsis AtKu80* gene demonstrates an essential role for AtKu80 protein in efficient repair of DNA double-strand breaks *in vivo*. *The Plant Journal* **31**: 517-528.

Wilson TE, Grawunder U, and Lieber MR. (1997) Yeast DNA ligase IV mediates non-homologous DNA end joining. *Nature* **388**: 495-498.

Wolf PG, Rowe CA, Sinclair RB, and Hasebe M. (2003) Complete nucleotide sequence of the chloroplast genome from a leptosporangiate fern, *Adiantum capillus-veneris* L. *DNA Research* **10**: 59-65.

Wyman C, and Kanaar R. (2006) DNA double-strand break repair: all's well that ends well. *Annual Review of Genetics* **40**: 363-383.

Xie Z, and Merchant S. (1996) The plastid-encoded *ccsA* gene is required for heme attachment to chloroplast *c*-type cytochromes. *The Journal of Biological Chemistry* **271**: 4632-4639.

Zhang Y, Hefferin ML, Chen L, Shim EY, Tseng H, Kwon Y, Sung P, Lee SE, and Tomkinson AE. (2007) Role of Dnl4-Lif1 in nonhomologous end-joining repair complex assembly and suppression of homologous recombination. *Nature Structural & Molecular Biology* **14**: 639-646.

Zoubenko OV, Allison LA, Svab Z, and Maliga P. (1994) Efficient targeting of foreign genes into the tobacco plastid genome. *Nucleic Acids Research* **22**: 3819-3824.

Zubko MK, Zubko EI, Zuilen KV, Meyer P, and Day A. (2004) Stable transformation of petunia plastids. *Transgenic Research* **13**: 523-530.

Zuo J, Niu Q, Moller SG, and Chua N. (2001) Chemical-regulated, site-specific DNA excision in transgenic plants. *Nature Biotechnology* **19**: 157-161.

VITA

Taegun Kwon earned a Bachelor of Science degree (1996) and a Master of Science degree (1998) in Agricultural Chemistry from Seoul National University, Seoul, Korea. He received an excellence in teaching award from the Seoul National University. He then worked in a pharmaceutical company, ShinDongBang Medex (1998-1999) and a non-profit research organization, Samsung Biomedical Research Institute (1999-2004) in Seoul, Korea. In 2004, Taegun attended Auburn University in Auburn, Alabama and received a Cell and Molecular Biology research fellowship. In 2005, he transferred to the University of Texas at Austin and joined David Herrin's laboratory to continue his doctoral work. Taegun has been a teaching assistant for several courses in botany, bacteriology, genetics and cell biology. He has supervised several undergraduate research students. Taegun is one of the few students to win both the Memorial Research Award (2008) and the Harold Bold Teaching Award (2010). He received a research fellowship (2008) from the Section of Molecular Cell and Developmental Biology. He was also awarded a Graduate School Fellowship (2008), a David Bruton fellowship (2009), and the C.P. "Pete" Oliver Memorial Endowment for Genetic Research Award from the Graduate School of the University of Texas at Austin. His publications include the following:

*Kwon T, Huq E, and Herrin DL. (2010) Microhomology-mediated and nonhomologous repair of a double-strand break in the chloroplast genome of *Arabidopsis*. *Proceedings of the National Academy of Sciences USA* **107**: 13954 -13959.

Lee BH, Nam HY, *Kwon T, Kim SJ, Kwon GY, Jeon HJ, Lim HJ, Lee WK, Park J, Ko JY, and Kim DJ. (2006) Paclitaxel-coated expanded polytetrafluoroethylene haemodialysis grafts inhibit neointimal hyperplasia in porcine model of graft stenosis. *Nephrology Dialysis Transplantation* **21**: 2432 -2438.

Chun J, Hyun S, *Kwon T, Lee EJ, Hong S, and Kang SS. (2005) The subcellular localization control of integrin linked kinase 1 through its protein-protein interaction with caveolin-1. *Cellular Signalling* **17**: 751-760.

Chun J, *Kwon T, Lee EJ, Hyun S, Hong S, and Kang SS. (2004) The subcellular localization of 3-phosphoinositide-dependent protein kinase is controlled by caveolin-1 binding. *Biochemical and Biophysical Research Communications* **326**: 136-146.

Chun J, *Kwon T, Lee EJ, Kim CH, Han YS, Hong S, Hyun S, and Kang SS. (2004) 14-3-3 Protein mediates phosphorylation of microtubule-associated protein tau by serum- and glucocorticoid-induced protein kinase 1. *Molecules and Cells* **18**: 360-368.

Lee YK, *Kwon T, Kim DJ, Huh W, Kim Y, Oh H, and Kawachi H. (2004) Ultrastructural study on nephrin expression in experimental puromycin aminonucleoside nephrosis. *Nephrology Dialysis Transplantation* **19**: 2981 -2986.

Chun J, *Kwon T, Kim DJ, Park I, Chung G, Lee EJ, Hong SK, Chang S, Kim HY, and Kang SS. (2003) Inhibition of mitogen-activated kinase kinase kinase 3 activity through phosphorylation by the serum- and glucocorticoid-induced kinase 1. *Journal of Biochemistry* **133**: 103 -108.

Lee HH, KIm DJ, Kim JA, Lee YK, Yoon SJ, Huh W, Kim YG, Oh HY, Kim Y, Shin SK, and *Kwon TG. (2003) Changes of the thyroid function in continuous ambulatory peritoneal dialysis patients. *Korean Journal of Nephrology* **22**: 96-101.

Lee YK, *Kwon TG, Kim DJ, Kim MK, Yoon SJ, Lee HH, Kim JA, Huh W, Kim YG, and Oh H. (2002) Expression of nephrin in puromycin aminonucleoside nephrosis in rat. *Korean Journal of Nephrology* **21**: 382-390.

Chun J, *Kwon T, Lee E, Suh P, Choi E, and Sun Kang S. (2002) The Na^+/H^+ exchanger regulatory factor 2 mediates phosphorylation of serum- and glucocorticoid-induced protein kinase 1 by 3-phosphoinositide-dependent protein kinase 1. *Biochemical and Biophysical Research Communications* **298**: 207-215.

Kim DK, Seo MY, Lim S, Kim S, Kim J, Carroll BJ, Kwon DY, *Kwon T, and Kang SS. (2001) Serum melanotransferrin, p97 as a biochemical marker of Alzheimer's disease. *Neuropsychopharmacology* **25**: 84-90.

Kwon T, *Kwon DY, Chun J, Kim JH, and Kang SS. (2000) Akt protein kinase inhibits Rac1-GTP binding through phosphorylation at serine 71 of Rac1. *Journal of Biological Chemistry* **275**: 423 -428.

Kim YH, *Kwon TK, Park S, Seo HS, Cheong J, Kim CH, Kim J, Lee JS, and Choi YD. (2000) Trehalose synthesis by sequential reactions of recombinant maltooligosyltrehalose synthase and maltooligosyltrehalose trehalohydrolase from *Brevibacterium helvolum*. *Applied and Environmental Microbiology* **66**: 4620-4624.

Kang SS, *Kwon T, Kwon DY, and Do SI. (1999) Akt protein kinase enhances human telomerase activity through phosphorylation of telomerase reverse transcriptase subunit. *Journal of Biological Chemistry* **274**: 13085 -13090.

Permanent address: 255-57 Waegwan 3-ri, Waegwan, Kyoungbuk 718800, Korea (South)

Permanent E-mail: tgkwon@utexas.edu

This dissertation was typed by the author.